

Exploring Virtual Rewards in Real Life:

A Gimmick or a Motivational Tool for Promoting Physical Activity?

Author: Yuvin Ha

Supervisor: Andrés Lucero

Advisor: Andrés Lucero, Maria Karyda

Master's Thesis
2019



Aalto University
School of Arts, Design
and Architecture

Author Yuvn Ha

Title of thesis Exploring Virtual Rewards in Real Life:
A Gimmick or a Motivational Tool for Promoting Physical Activity?

Department Department of Design

Degree programme Collaborative and Industrial Design

Year 2019

Number of pages 72

Language English

Abstract

Virtual rewards, digital badges or points, are often awarded to extrinsically reinforce desirable user behaviors. Acknowledging their motivational power for increased user engagement and enhanced playfulness, virtual reward systems are widely employed in contemporary activity tracking devices and services. While the characteristics of the current virtual reward systems vary and previous studies point to mixed outcomes around the effectiveness of virtual rewards for encouraging physical activity, there is little empirical research on the actual use and impact of virtual rewards provided from activity tracking tools in a real-life environment. To address the knowledge gaps on the use and effects of virtual rewards, I investigate the lived experiences of users with the virtual reward system of the activity tracking tool and their effects and value in the physical activity practices of individuals. To understand this both qualitatively and quantitatively, I conduct user experience research using mixed methods. *Diary-based autoethnography* research is preceded using three popular activity trackers, thereby having and documenting situated experiences with two different digital badge systems and one points-based system. My autoethnographic exploration is followed by an online survey with 113 existing users of activity tracking devices. The collected data is analysed through content and statistical analyses. By interpreting the findings from both studies and the reviewed literature, this thesis illustrates in great detail how virtual rewards interact with users and other system features in real-life practices. In addition, it provides a clearer picture of the effects of virtual rewards on people's engagement in physical exercise and what makes some of them more valuable. Finally, it offers a set of considerations for the design of virtual reward systems that can lead to more user-centric, multi-layered, and meaningful virtual reward experience.

Keywords Activity Tracker; Physical Activity; Virtual Reward; Digital Badge; Point; Autoethnography; User Experience

Acknowledgments

I would like to first thank my supervisor Andrés Lucero and advisor Karyda Maria for taking their valuable time and giving me constructive advice and suggestions. Without their help, this work would never have been possible.

I also thank Antti Salovaara, who gave me much valuable advice on the statistical analysis of the survey results. He led me in the right the direction when I had a hard time handling quantitative data.

Also, I gratefully acknowledge the help provided by Katri Lassila, a workshop master in Väre Takeout. With her support, I was able to conduct autoethnographic research using three different activity trackers from major companies with limited funds.

In addition, this research was benefited from Department of Design Grant. Using the grant for recruiting the survey participants, I was able to reach 113 activity tracker users and collect their invaluable experiences and voices.

Lastly, I am grateful to the survey participants. Without their genuine and thoughtful answers, I could not have provided rich insights into the lived user experience of virtual rewards.

Table of Contents

1. Introduction.....	1
1.1. Background	2
1.1.1. Prevalence of Activity Tracking Devices	2
1.1.2. Activity Trackers	2
1.1.3. Virtual Rewards as Additional Motivator.....	3
1.1.4. Potential and Risk of Virtual Rewards	3
1.2. Research Aim, Objectives and Questions.....	5
1.3. Research Approach and Methods	6
2. Literature Review	7
2.1. Virtual Rewards as Game Element	8
2.1.1. Meaningful Gamification.....	8
2.1.2. Digital Badges and Points as Virtual Rewards	8
2.1.3. Application Cases of Badges and Points in Reward System.....	10
2.2. Virtual Rewards as Extrinsic Motivator	12
2.2.1. Activity Tracking and Virtual reward Experience in Real Life.....	12
2.2.2. Approaches to Extrinsic Rewards for Increased Physical Activity.....	14
3. Desk Research	16
3.1. Method and Procedure	17
3.2. Current Designs of Virtual Reward System	18
3.2.1. Activity app from Apple.....	18
3.2.2. Fitbit app from Fitbit.....	19
3.2.3. Garmin Connect app from Garmin	20
3.2.4. Samsung Health app from Samsung	21
3.2.5. Google Fit app from Google.....	22
3.3. Comparative Analysis.....	23
4. User Studies: Autoethnographic Diary Study.....	25
4.1. Methods.....	26
4.1.1. Research Materials.....	26
4.1.2. Data Collection	26

4.1.3. Data Analysis.....	28
4.2. Findings.....	29
4.2.1. Content.....	29
4.2.2. Aesthetic.....	32
4.2.3. Access	34
4.2.4. Trust.....	34
4.2.5. Communication	36
4.2.6. Goal-setting	37
4.4. Summary of Findings	39
5. User Studies: Online Survey.....	40
5.1. Methods	41
5.1.1. Participants.....	41
5.1.2. Questionnaire.....	41
5.1.3. Data Analysis.....	44
5.2. Results.....	45
5.2.1. Participants Characterization.....	45
5.2.2. Fitness Tracking Practices	45
5.2.3. Experiences with Virtual Rewards	47
5.3. Summary of Results	61
6. Discussion.....	62
6.1. Key Findings.....	63
6.2. Design Considerations	66
6.2.1. User-Centered Reward System	66
6.2.2. Multi-Layered Virtual Reward Experience	67
6.2.3. Meaningful Reward Experience.....	68
6.3. Limitations	70
7. Conclusion.....	71
References.....	73

List of Tables

Table 1	Activity tracking app and reward type by company.....	17
Table 2	Comparison between virtual reward systems in activity trackers of companies.....	24
Table 3	A list of Likert scale items.....	43
Table 4	Frequencies of Familiarity of the current badge/points system.....	48
Table 5	Frequencies of Satisfaction of the current badge/points system.....	48
Table 6	Familiarity and satisfaction between age groups.....	48
Table 7	Familiarity and satisfaction between usage groups.....	48
Table 8	Code frequencies of 16 categories identified.....	50
Table 9	Mean, standard deviation, and percentage of frequencies of Likert scale items	56
Table 10	Final constructs after factor analysis.....	56
Table 11	Mean and standard deviation of the mean scores of System Quality and Perception and Effect by virtual reward system.....	57
Table 12	One-way Analysis of Variance of system Quality and Perception and Effect by virtual reward system	58
Table 13	Welch's t-test for System Quality and Perceived Effects (Younger vs. Older)	58
Table 14	Independent Samples t-test for System Quality and Perceived Effects (Long-term vs. Short-term).....	58
Table 15	Correlation between satisfaction and individual Likert scale items	59
Table 16	Correlation between familiarity and individual Likert scale items.....	59

List of Figures

Figure 1 Research process	6
Figure 2 Screen images of the Activity app.....	18
Figure 3 Screen images of the Fitbit app	19
Figure 4 Screen images of the Garmin Connect app	20
Figure 5 Screen images of the Samsung Health app	21
Figure 6 Screen images of new Google Fit app.....	22
Figure 7 Images of research materials.....	26
Figure 8 Timeline with an overview of autoethnographic research	28
Figure 9 Collected data from autoethnography	28
Figure 10 Result of affinity diagramming with 6 major themes.....	29
Figure 11 Overall flow of the survey.....	42

01

Introduction

1.1. Background

1.1.1. Prevalence of Activity Tracking Devices

Personal informatics, a class of tools that support people to gather and reflect on personal information (Li, Dey, & Forlizzi, 2010), have become more prevalent in various aspects of everyday life, from physical activity to medication and finance. In particular, self-tracking tools for physical activity are getting widely adopted not only for a specific purpose such as preventing a health-related problem, but also to satisfy curiosity about their functions or potential impacts (Jarrahi, Gafnowitz, & Shin, 2018), as people's interests in well-being and healthier lifestyles have been increasing. In response to the growing needs for tracking physical activity, many affordable products are available in the market in the form of wearable devices, which are normally referred to as an activity tracker. The IDC (2018a), International Data Corporation, forecasted that the global wearable market will continue to grow, mostly fuelled by wrist-worn activity trackers and smartwatches, expecting the annual total shipments to reach 190.4 million units in 2022 from 122.6 million units in 2018 with a compound annual growth rate (CAGR) of 11.6% in the next five years.

1.1.2. Activity Trackers

Activity trackers, also known as fitness trackers, are commercialized tools capable of tracking and monitoring physical activity and other physiological information by automatically collecting a variety of data (Shih, Han, Poole, Rosson, & Carroll, 2015). With multiple embedded sensors such as accelerometers, altimeters, and even electrocardiogram, people are enabled to monitor activity data such as step counts, distance moved, speed and pace, calories burned, heart rate, hours slept and so on (Shih et al., 2015). Also, they are often accompanied with mobile applications and online servers where all data are stored, shared with other users or friends, and visualized to show the progress and the feedback to their users. These technical advances in commercialized activity tracking devices can benefit people by promoting physical activity with the numerical feedback (Fritz, Huang, Murphy, & Zimmermann, 2014), helping manage the overall health more effectively with less effort (Montgomery-Downs, Insana, & Bond, 2012), and leading to a positive behavioral change in health and well-being. There are a couple of different types available in the market, but wrist-worn activity trackers are the most widely used. In a broad sense, smartwatches also can be considered as one type of activity tracking devices in that they provide their users with almost the same functionality. For these reasons, wrist-worn activity trackers and smartwatches are going to be dealt with in this study.

1.1.3. Virtual Rewards as Additional Motivator

Persuasive technology is the term to describe interactive computing systems which are designed to change attitudes or behaviors of the users (Fogg, 2003). Activity trackers are inherently built on *self-monitoring* technology, one of seven types of persuasive technology tools identified by Fogg (2003). While *self-monitoring* allows people to have more awareness about their behaviors and themselves by presenting ongoing data such as step counts, flights climbed, or heart rates, it is also asserted that the information provided is not motivational enough to encourage a sustained engagement in a healthier lifestyle for a majority of people (Waltz, 2012). In line with this, persuasive strategies are often implemented in combination, and along with *self-monitoring*, one of the techniques used most in activity trackers is *reinforcing* (Fritz et al., 2014; Karapanos, Gouveia, Hassenzahl, & Forlizzi, 2016; Orji, Vassileva, & Mandryk, 2014). It is often approached by establishing a system of virtual rewards, where users are extrinsically reinforced for desirable behaviors via digital incentives mostly in the form of digital badges and points. To harness its motivational power for increased user engagement and enhanced playfulness, leading companies in this industry including Apple, Fitbit, Google, and Samsung have introduced their own virtual reward systems to the products and services. However, the approaches to designing a virtual reward system taken by the companies vary, and many characteristics of a virtual reward system, such as the forms of reward and reward mechanisms, are different among each other. Though these differences may yield different behavioral and motivational outcomes, the actual use and impact of the current designs have been underexplored, and the mixed views on the digital rewards offered by activity tracking devices have been only partially reported by studies on the use of activity trackers in everyday life. (Fritz et al., 2014; Jarrahi et al., 2018; Kappen, Mirza-Babaei, & Nacke, 2017; Vooris, Blaszk, & Purrington, 2019).

1.1.4. Potential and Risk of Virtual Rewards

Meanwhile, there has been some concern about the detrimental effect of virtual rewards on intrinsic motivations since virtual rewards are essentially virtual forms of extrinsic rewards. The study of the impacts of externally mediated rewards on internal sources of motivation has a long history in the field of social psychology. Deci, Koestner, and Ryan (1999) indicated through a meta-analysis of 128 related studies that the use of extrinsic rewards involves a risk of diminishing people's intrinsic motivation, caused mostly by tangible rewards. However, it is not clear yet that the same issue can arise in terms of intangible form of rewards. Mekler, Brühlmann, Opwis, and Tuch (2013) presented a contradictory result through their experiment in an online image annotation task, showing that points did not undermine participants intrinsic motivation, but rather increased. Zichermann and Cunningham (2011) held that well-designed extrinsic rewards in gamified systems can trigger intrinsic motivation. With regard to this, careful implementation of extrinsic reward systems is required to trigger

internal motivation and maintain it (Richter, Raban, & Rafaeli, 2015), and the success of implemented systems can be predicted by the context of being used and the qualities of users (Hamari, Koivisto, & Sarsa, 2014). Consequently, a more nuanced understanding of users and their situated context is fundamental for the use of virtual rewards in order to yield intended outcomes while minimizing the adverse effect that external incentives may bring about while taking advantage of the potential opportunities of digital rewards.

1.2. Research Aim, Objectives and Questions

Despite a need for better knowledge on users and the actual context of use to create a positive, meaningful results through digital rewards, we still have a limited understanding of people's lived experiences with virtual reward systems of activity tracking tools. Also, it is still not clear what effect virtual rewards have on user motivation and how they influence people's engagement in physical exercise. To address the knowledge gap described above, this study pursues to answer the following research questions:

- How do people experience different virtual rewards provided by activity tracking tools in everyday practice?
- What effects and value virtual rewards have on people's behaviors and motivations for physical activity?

In other words, I aim to investigate people's real-life experiences and perceived effects of the virtual reward systems in activity tracking devices, thereby exploring desirable ways of a virtual reward system for promoting physical activity. Following research objectives would facilitate achieving this aim:

- To illustrate how virtual rewards have been incorporated into real-life experiences of wearable activity tracker users
- To examine actual users' perceived effects and value of, and their expectations for virtual rewards provided by activity tracking tools
- To provide the implications for designing a virtual reward system aimed to promote physical activity

1.3. Research Approach and Methods

To understand, both qualitatively and quantitatively, how virtual rewards are engaged with actual users in real-life contexts and what influence they have on users' motivation and behaviors, I conduct user experience research using mixed methods: a *diary-based autoethnography* and an *online survey*. This mixed approach can be justified by the previous literature. Klasnja, Consolvo, and Pratt (2011) pointed out the importance of field studies to understand the impact of technologies for health behavior change and how they interact with users and other situated factors, which is generally hard to be achieved by a traditional way of HCI (Human-computer interaction) research such as deploying a research prototype. Autoethnography is one way to conduct field research using a first-person point of view. To be more specific, autoethnography, derived from qualitative research in the social sciences, is a form of autobiographical research in which the researcher's own thoughts and lived experiences as the central element of the study are expressed and systematically analysed to understand social, cultural phenomenon (Ellis, Adams, & Bochner, 2011; Reeves, Kuper, & Hodges, 2008). Conducting autoethnography allows the researchers to gain first-hand user experience of the device, by which generates deep and rich insights (Cecchinato, Cox, & Bird, 2017). Such merits, which is barely possible through other research methods, have made first-person research more important in HCI research as the private areas of life, a space where situations are intimate and long-term investigation is required such as home and wearables, is becoming the research topics in the HCI community (Desjardins & Ball, 2018). However, its inherent nature—subjectivity, emotionality, and researcher's influence on research—is often challenged (Lucero, 2018). In this sense, conducting a third-person perspective research along with autoethnography can be expected to develop a nuanced understanding and elicit novel insights while alleviating the challenges that autoethnographic research faces. For these reasons, I decided to approach the research questions through a combination of a diary-based autoethnography and an online survey.

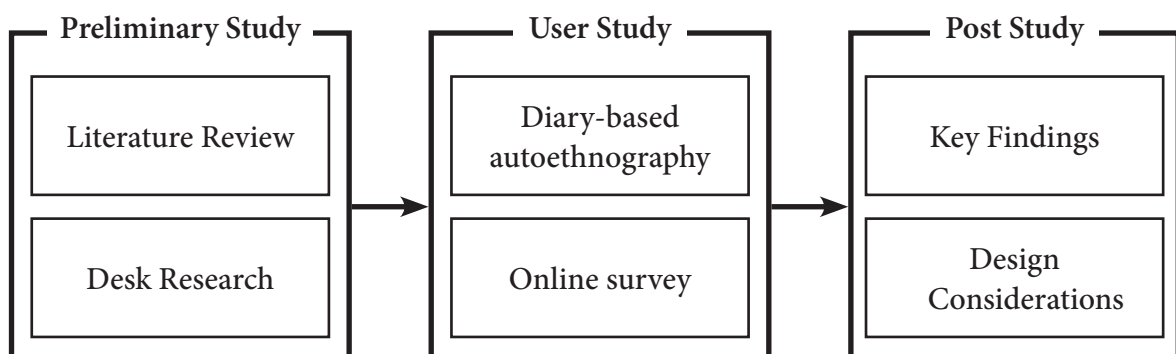


Figure 1. Research process

02

Literature Review

Virtual rewards are digital or intangible incentives awarded for desired behaviors or outcomes, given in the form of points, badges, or extra game items or resources (Zuckerman & Gal-Oz, 2014). A system of virtual rewards is commonly included in games as a fundamental mechanism to increase player engagement, direct gameplay behaviors, and promote enjoyment (Goh, Pe-Than, & Lee, 2017). Thus, virtual reward systems in activity tracking devices can be identified as the result of gamification, which is the integration of game design elements in non-game contexts (Deterding, Dixon, Khaled, & Nacke, 2011). Therefore, literature about gamification and game reward system is worth examining in this research and the most relevant work is introduced at first. Then, I present previous research on everyday uses of activity trackers to grasp the current state of knowledge in this field. In this process, the available evidence on the effectiveness of current virtual reward systems in activity trackers are reviewed in detail. Lastly, I examine prior studies in the context of digital health interventions for increased activity level. Researchers in the field of Human-computer interaction (HCI) have developed persuasive systems for promoting physical activity over the last decade, in which various ways to extrinsically reward users have been proposed and evaluated. Reviewing these works is worthwhile in that they can give insights into how people are influenced by different reward systems within this specific context.

2.1. Virtual Rewards as Game Element

2.1.1. Meaningful Gamification

Gamification emerged in an attempt to leverage aspects of play to make non-game activities more fun and engaging. Gamification has been receiving huge attention from practitioners and researchers in recent years, and its application has been extended to diverse areas, from education to health and commerce. In the meantime, criticism arose among some scholars from the hype around gamification. Deterding (2013) argued that existing gamified systems are 1) not systemic by merely adding game design elements, 2) reward-oriented by focusing on motivating through extrinsic rewards, 3) not user-centric by emphasizing the goals of the system owner, and 4) pattern-bound by limiting themselves to a small set of feedback interface design patterns. Cugelman (2013) pointed out the misconception of gamification, claiming that gamified systems work effectively only when used in the right way based on a thorough understanding of game tactics. In line with this movement, efforts have been made to establish a theoretical foundation for more effective gamification. Nicholson (2012) introduced the concept of meaningful gamification by exploring the theories of motivation, learning, and user-centered design, and defined meaningful gamification as “the integration of user-centered game design elements in non-game contexts” (p. 5). While pointing out that the mere use of game elements without a careful consideration of the underlying non-game setting and users will result in a poor gamification experience, he claimed that gamified systems are required to help people find meaningful connections between game elements, activity, and their own goals and desires by providing relevant information to users’ interests, allowing user customization, and increasing the transparency of the system. Hamari et al. (2014) examined 24 peer-reviewed empirical studies of gamification based on a proposed conceptual framework for gamification which consists of motivational affordances, psychological outcomes, and further behavioral outcomes. They indicated that the effectiveness of gamification can become different depending on the context being gamified and user types while gamification mostly brings positive effects on users. This result is also aligned with a Deterding (2011)’s theoretical study on situated motivational affordances. Also, Vassileva (2012) described the importance of modelling users and adapting the motivational strategies to their interests and needs. Therefore, understanding those who actually use the systems in real life and the context where the systems are employed is a key for gamified systems to produce intended motivational and behavioral outcomes.

2.1.2. Digital Badges and Points as Virtual Rewards

Virtual reward systems are found as the most common manifestation of gamification in various contexts, and virtual rewards are often awarded to users through digital badges or

points in gamified systems. They are expected to function in the same way as they do in games by building the relationship between user actions and outcomes, giving a sense of achievement and pleasure, and allowing users to reflect their status and collections (Wang & Sun, 2011). However, there are also some differences between badges and points in terms of how information is represented (non-numeric vs. numeric), where values come from (social vs. individual), and their assumed roles, which could yield different effects on users' perceptions and behaviors (Goh et al., 2017). Thus, a proper review of both systems is important even though the comparison between the two systems is not the main purpose of this study. For this reason, fundamental knowledge of digital badges and points are described first, and it is followed by prior application cases of both types of rewards in gamified systems with their findings and limits.

Firstly, a digital badge is a digitized image or an icon that contains associated information or metadata (Shields & Chugh, 2017) indicating one's accomplishment, skill, quality, or interest (HASTAC: Humanities, Arts, Science, and Technology Alliance and Collaboratory, n.d.). It is also regarded as one of communication interfaces between users and systems, in which badges encapsulating information specified by the designer are transferred to users when certain requirements are met (Fanfarelli & McDaniel, 2015). It is theoretically constituted with three primary elements, which are signifier, completion logic, and reward (Hamari & Eranti, 2011). To be specific, signifier is a visible part for users, consisting of a name, a visual component such as an icon or a badge, and a description of the requirement and consequence of action. Next, completion logic is the foundational logic of a badge, which defines pre-requirements, trigger, conditions, and the amount of times the trigger has to be activated under the conditions. Lastly, reward is the compensation for users after unlocking the badge, and can be categorized into three, which are in-game, achievement game, and out-game. While badges are primarily used as extrinsic rewards in digital systems, scholars suggested different roles that digital badges could take on in various contexts. In an attempt to create a taxonomy of achievement systems in games (Montola, Nummenmaa, Lucero, Boberg, & Korhonen, 2009), they identified digital badges as sub-goals in a secondary reward system that add additional motivations for playing the core game. In an educational context, Jobe (2014) stated that badges indicate recognition, validation, and accreditation of non-formal learning by displaying to the public the skills or accomplishments that people have learned or completed. Antin & Churchill (2011) presented five individual and social functions for badges in social systems such as Foursquare, StackOverflow, and Wikipedia: goal-setting, instruction, reputation, status/affirmation, and group identification. With a focus on their social function, badges also can represent social norms of a system by manifesting the types of valuable activities and interactions (Antin & Churchill, 2011; Cugelman, 2013).

While badges represent people's interests and experiences in a graphical way, points are numerical scores in a system that are awarded for users' behaviors (Goh et al., 2017). In

other words, Point systems use numbers to signify a user's performance (Wang & Sun, 2011). Sometimes, a composite metric is created and used in order to reduce the complexity of data into a single number (Zichermann & Cunningham, 2011). In the context of this study, NikeFuel is a good example of this. It is a universal point system developed by one of the athletics companies, Nike, to allow its users to measure all activities that they engaged in everyday life and see them in a single, abstract number. Zichermann and Cunningham (2011) listed five different kinds of points that can be used in gamified systems: experience points, redeemable points, skill points, karma points, and reputation points. Since the latter two are more complex and rarely used in classic games, the details of the first three types are introduced in this study. Experience points accrue from every single activity a user does within the system. They do not have any monetary value in the system, so that they cannot be redeemed. While skill points share some common traits with experience points, they serve as more like bonus points allowing users to earn this point from specific activities while they enjoy the core game. On the contrary to the previous two types of points, redeemable points are allowed to be exchanged into some other kind of reward, so it establishes the base of a virtual economy. When it comes to the roles that points play in the system, points primarily provide feedback on users' actions and performance (Lewis, Swartz, & Lyons, 2016; Mekler et al., 2013; Richter et al., 2015). In line with this, points can perform as tools for self-assessment and comparison (Wang & Sun, 2011). Also, points indicate an individual's progression, and can be used as a goal-setting tool (Goh et al., 2017; Richter et al., 2015).

2.1.3. Application Cases of Badges and Points in Reward System

Though the interest in gamification was rapidly increased and many digital services in diverse areas that are gamified with digital badges and points were available to customers, the effectiveness of gamification approach remained in question. Along with theoretical consideration on this matter taken by some scholars (Cugelman, 2013; Hamari et al., 2014; Nicholson, 2012; Vassileva, 2012), other scholars took more experimental approaches to clarify the potential of gamification in their field of study. Montola et al. (2009) implemented a point-based rewards system to a location-based photo sharing service for increased user participation, in which users were incentivized through points for four different tasks. A two-month field trial with 20 participants showed that many participants did not appreciate the added value that the points system can deliver while a few found some hedonic value from it. In contrast, Mekler et al. (2013)'s examination of the impacts of points and a meaningful framing on internal motivation and performance in an online image annotation task indicated that points system predicts the quantity of outcomes that participants generate while a meaningful frame leads to higher quality of work. Also, they reported that both points and a meaningful framing did not diminish users' intrinsic motivation, but rather increased to a similar degree. When it comes to the health-related interventions, Zuckerman

and Gal-Oz (2014) developed three different types of research prototypes and tested them with participants in order to compare the effectiveness of different game design elements to promote opportunistic physical activity. In their field studies, most participants perceived a points system as meaningless while being more interested in actually measured data rather than an arbitrary number. Moreover, they did not find any significant difference between gamified versions and the quantified version in terms of the level of physical activity that participants engaged in. Munson and Consolvo (2012) developed a mobile application and conducted a four-week field study to explore how the strategies of goal-setting, rewards, self-monitoring, and sharing encourage physical activity. In their study, digital rewards analogous to badges failed to motivate most participants although the reactions to virtual rewards were positive in their preliminary survey, posing a question about how such rewards should be designed within this context. In sum, prior studies on the effectiveness of virtual rewards for promoting desirable behaviors yielded mixed results. The rewards were persuasive in certain contexts, whereas they were not considered as valuable and did not affect users' attitude and performance in other contexts. In some of the reviewed studies, researchers suggested that the behavioral and motivational outcomes of virtual reward system may vary depending on the context being employed, the activities appreciated by the system through awarding virtual rewards, and the way of implementation.

2.2. Virtual Rewards as Extrinsic Motivator

In the previous section, I considered virtual rewards as game elements, reviewing the relevant literature about gamification, game design elements, and their application cases in various fields of study with a focus on two types of the most commonly used virtual rewards, digital badges and points. In the reviewed studies, researchers pointed out that it is essential to take account of the context of use, the activities that are gamified, and the qualities of users in order for gamified systems to yield intended effects. Fortunately, the widespread adoption of activity trackers allows us to deepen our understanding of actual users, their lived experiences, and the barriers that they face. In an effort to do this, some experience of, and opinions on the virtual rewards systems in current activity trackers are also uncovered. Along with the review of these studies and their findings, prior attempts to externally promote physical activity are also presented in this section. Examining these trials could provide further consideration of how to support people through extrinsic motivators.

2.2.1. Activity Tracking and Virtual reward Experience in Real Life

Since there has been a rapid proliferation of consumer products for tracking health- and fitness-related activities over the past few years, research on the real experience and the actual impact of activity tracking devices in everyday life is still in its developing stages. The recent studies pointed out that most tools have failed to sustain a long-term engagement with a majority of users. The Endeavour Partners (2014), a management consulting firm, revealed that one-third of activity tracker users in the United States (US) stopped using their tracking devices within six months since they had started using them. Moreover, Jakicic et al. (2016) indicated in their longitudinal study on the effect of wearable technology in a weight loss intervention that activity trackers failed to offer the long-term advantage in improving one's health. Users' abandonment of activity trackers also raised academic attention in the HCI community, followed by further research aiming to develop more effective strategies to ensure sustained engagement by understanding motivations, actual use and impact, and problems that hinder a long-term use of activity tracking devices. Gouveia, Karapanos, and Hassenzahl (2015) studied how users' engagement with activity trackers changes over time in real life by developing an activity tracking app named 'Habito' and analysing the data from 256 users. Another attempt was made by Fritz et al. (2014). They examined people's changing perspectives and perceived impacts on their activities and attitudes over long period of use through in-depth interviews with 30 participants who had been using activity trackers for between 3 months and 54 months. More recently, Vooris et al. (2019) explored the reasons for purchasing activity trackers and their usage patterns through focus-group interviews with 31 participants. The results of these studies indicated that using activity tracking devices

bring positive changes in people's routines or behaviors although the use of activity trackers may not yield direct effects on the improvement of physical or mental health (Fritz et al., 2014; Vooris et al., 2019). However, those impacts can be varied depending on people's commitment to exercise, and current activity tracking devices work more effectively for those who have the intention for change but do not know how (Gouveia et al., 2015). Also, researchers pointed out the limited support of current trackers for long-term use while users' activity goals and roles of activity trackers in everyday practice have changed over time (Fritz et al., 2014; Gouveia et al., 2015). They stated that the current systems fail to reflect users' changing informational needs and interests, which leads to users' disengagement with activity tracking devices. In addition, barriers to sustained engagement with activity trackers were also examined (Harrison, Marshall, Bianchi-Berthouze, & Bird, 2015). In their surveys and interviews with 24 participants including 8 participants who abandoned tracking, researchers argued that providing the accuracy of tracking, appropriate credits for users' efforts, and social functionality are crucial for long-term engagement. Meanwhile, some evidence on the effects of virtual rewards in current activity trackers were found in these studies. Fritz et al. (2014)'s investigation showed that system rewards like badges and points had a positive effect on more than half of participants by promoting more physical activity, giving a sense of achievement, and serving as an alternative goal. Positive responses to virtual rewards were also given by Vooris et al. (2019)'s examination of how the gamification elements impact on users' usage patterns of activity trackers, indicating that badges help people stay active, let them move more to earn them, and check the app more often. These findings are in conflict with the results from Munson and Consolvo (2012)'s experiments. Other studies pointed to more mixed outcomes. Jarrahi et al. (2018) carried out a qualitative study based on in-depth interviews with users of Fitbit activity trackers to explore how different types of prior motivations affect the perception and adoption of such devices. They suggested the typology of users, in which users are grouped into five based on the individuals' activity levels, pre-existing motivations, and attitudes toward their personal data. Their findings showed that motivational features such as virtual rewards are effective only for certain types of users who have already wanted to be more active. Kappen et al. (2017) paid their attention to differences in the effects and preferences of motivational elements in activity tracking devices between age groups. By exploring the motivational techniques for promoting physical activity, they identified that badges were beneficial to the age group between 18 and 29 while they failed to create a meaningful value or even brought a negative impact on motivation for engaging in physical activity in other age groups. On the other hand, Novak and Loy (2018) discussed the limited motivational capability of the virtual reward system that one of the authors experienced from the first-person research with a Garmin activity tracker. They described that the current reward system fails to provide any real context or meaning for users and to fit the advancement of personal goals. As described above, the results of the previous studies were mixed regarding the effectiveness of digital rewards provided by current

activity tracking devices. This could have been caused by a limited number of participants in qualitative research. In most cases, the findings of the studies were drawn from the in-depth interviews or focus groups with a small number of users, in which there is always a risk that those who were recruited for research may not represent the entire group of users (Fritz et al., 2014; Vooris et al., 2019). With this regard, a need for research with a larger and broader population was came up by a few researchers (Jarrahi et al., 2018; Vooris et al., 2019). Moreover, it was confirmed that most studies only focused on motivational effects of virtual rewards, providing a limited understanding of how virtual rewards have been incorporated into real-life experiences of activity tracker users.

2.2.2. Approaches to Extrinsic Rewards for Increased Physical Activity

Virtual rewards are inherently a digital form of extrinsic rewards, and digital badges and points are one of the most commonly used types of system rewards in health-related technologies. As the social needs for digital health interventions to address increased sedentary lifestyles have been growing, many researchers also have experimented with different approaches to the system rewards for increased physical activity. One approach taken by scholars is incorporating real-world physical activity into the virtual games which are developed for this purpose. Fish'n'Steps (Lin, Mamykina, Lindtner, Delajoux, & Strub, 2006) is a social computer game designed for encouraging people to take more step counts, in which users' daily progress towards their step goals are reflected to the growth and the emotional state of their own virtual fish in both positive and negative ways. The results of a 14-week study with 19 participants showed the positive influence of the intervention on two thirds of the participants in terms of activity level. The authors also confirmed that the negative reinforcement may lead people to stop using the system within this context. For the same purpose, UbiFit Garden (Consolvo et al., 2008) was also developed. The system used the background screen on the mobile phone, displaying the virtual garden which is filled with butterflies and flowers based on the level of physical activity and goal attainment. In a 3-week field trial with 12 participants, features of the system were favourably received, especially the glanceable display was regarded as crucial. While the first two cases visualized physical activity levels via virtual creatures in the game world, Berkovsky, Coombe, Freyne, Bhandari, and Baghaei (2010) used a direct game related motivator. They applied the PLAY, MATE! design to a public game, Neverball, in which users were encouraged physical activity to get time-based rewards. The more extra time they earn, the higher their chances are of succeeding in the game. A user study with 180 participants proved that the proposed game was effective to motivate people to perform more physical activity and did not have a harmful effect on perceived enjoyment of playing. More recently, some researchers have investigated ways to explicitly reward people not in the virtual environment, but in the

physical world. Khot (2013) built Sweat-Atoms, a 3D modeling and printing system, that uses heartbeats of individuals to generate abstract 3D designs. Although the sample size was small ($N=5$), the participants in their preliminary study were favorable to the system, seeing it as a new way of enjoying exercises. Taking a similar approach, Stusak, Tabard, Sauka, Khot, and Butz (2014) designed Activity Sculptures, a series of physical visualizations of tracked running data, and evaluated with 14 participants. In a 3-week field study, participants were asked to track their running data with a mobile application and received the unique sculptures created based on their data within 3 days after submission. The study confirmed the positive influence of material representation of tracked data on motivation, self-reflection, and social interactions. Moreover, Novak and Loy (2018) developed two tangible reward systems by utilizing 3d printing technology and the Internet of things based on their understanding of the limited motivational capability of the current virtual rewards from first-person research. The authors attempted to create a more personalized and meaningful experience of getting and seeing rewards beyond digital interface, aiming to have younger generations sustain a longer engagement in physical activity. While some studies showed the effectiveness of their systems with a small number of participants, the limited nature of the physical representation was also revealed. Stusak et al. (2014) reported that costs of materials and time-consuming printing process need to be addressed to implement this sort of system in the real world. In this regard, virtual rewards are more feasible since they are cost effective (Montola et al., 2009) and can be provided immediately or delayed for greater rewards (Wang & Sun, 2011).

Desk Research

Before carrying out user studies, it is essential to have sufficient knowledge of the current designs that are publicly available at present. I conducted a review of virtual reward systems in existing activity tracking devices and services provided by major companies in the wearable market to gain a broad understanding of what kind of digital rewards are offered, how they are employed in the tracking systems, and how users can earn and access their achievements.

3.1. Method and Procedure

According IDC's *Worldwide Quarterly Wearables Tracker* reports in 2018 (IDC, 2018b) and 2019 (IDC, 2019), top 5 companies of the wearable market in 2019 were Xiaomi, Apple, Huawei, Fitbit, and Samsung while those in 2018 were Apple, Xiaomi, Fitbit, Huawei, and Garmin. Those top 5 companies accounted for 53.4 percent and 65.7 percent of market share in 2018 and 2019, respectively. Thus, I examined first the products and services provided by the companies which were ranked in either report: Apple, Fitbit, Garmin, Huawei, Samsung, and Xiaomi. Since each company provides the dedicated mobile application for its wearable products, I downloaded those applications from the US version of Apple App Store. The reviewed mobile applications are listed in Table 1.

Table 1

Activity tracking app and reward type by company

Companies	Types of Activity tracking Devices	Mobile Applications	Types of Virtual Rewards
Apple	Smartwatches	Activity app	Digital badges
Fitbit	Smartwatches/Activity trackers	Fitbit app	Digital badges
Garmin	Smartwatches/Activity trackers	Garmin Connect app	Digital badges
Huawei	Smartwatches/Activity trackers	Huawei Health app	Not confirmed
Samsung	Smartwatches/Activity trackers	Samsung Health app	Digital badges
Xiaomi	Activity trackers	Mi Fit app	Not confirmed

By browsing their official websites and reviewing mobile applications, whether to adopt a virtual reward system in the company's activity tracking system and the types of reward were found (Table 1). While Huawei and Xiaomi did not implement a virtual reward system in their devices and services, other companies employed their own digital badge systems. These four digital badge systems, therefore, were further analysed. In addition to these four, I paid extra attention to the Google Fit app, which is a health and activity tracking platform developed by Google and first released in 2014. Google redesigned the app in 2018, through which new point-based reward system was introduced. Given that other major companies are awarding digital badges to their users as shown in Table 1, Google's recent attempt is worth investigating further to understand how different forms of virtual rewards affect people's attitudes and behaviors regarding physical activity. In the following section, the virtual reward systems embedded in activity tracking platforms of the five companies mentioned above are briefly reviewed.

3.2. Current Designs of Virtual Reward System

3.2.1. Activity app from Apple

Activity app is one of Apple Watch's default apps for activity tracking, and exclusively available to Apple Watch users. It allows users to track how much they move, how many minutes they exercise, and how often they stand up in their daily lives by representing them with three colored rings named the Move, Exercise, and Stand ring, respectively. Users are given the daily goals for those three rings while the goal of Move ring is only allowed for users to change on their own. Users are also enabled to check their progress of daily goals on the Apple Watch as well as the mobile version of the Activity app on the iPhone. When it comes to virtual rewards, the app awards several kinds of digital badges to its users who satisfy the predefined conditions. Badges are largely divided into three categories: standard badges, social badges, and time-limited challenges. The requirements for badges are defined based on the attainment of daily goals—the rate of goal achievement or the number of times the goals are achieved in a row as well as in total. Furthermore, the app distributes badges when users engage in various types of workout for the first time or beat their best records. Users are allowed to see the list of badges that are both already earned and available in the future with their progression on their smartphones, which became available on the Apple Watch as well with the introduction of watchOS 6, an operating system for Apple Watch.



Figure 2. Screen images of the Activity app: History tab (left), Awards tab (right), and detail of a badge (right)

3.2.2. Fitbit app from Fitbit

Alongside the activity trackers and smartwatches, Fitbit offers the Fitbit app for various versions of mobile operating system and individual's web dashboard on its official website. The app and web dashboard provide almost the same functions, enabling users to track activity, sleep, and weight, record workouts, and connect with friends, family and other users. As with Apple Watch, people can monitor the tracked data on their wrist in real time. Moreover, users are allowed to manually modify daily activity goals, such as steps, calories burned, and distance, on their mobile app. Unlike Apple Watch, digital badges offered by Fitbit have nothing to do with the completion of daily goals. There are several thresholds for digital badges in terms of steps, floors climbed, and weight, and users receive them once they hit those thresholds during a day or in a lifetime. In addition to badges, users are able to unlock the trophies by participating in challenges, in which they compete and compare with friends and family. Users are able to see the collection of the badges that they have achieved on the mobile app as well as on the web dashboard.

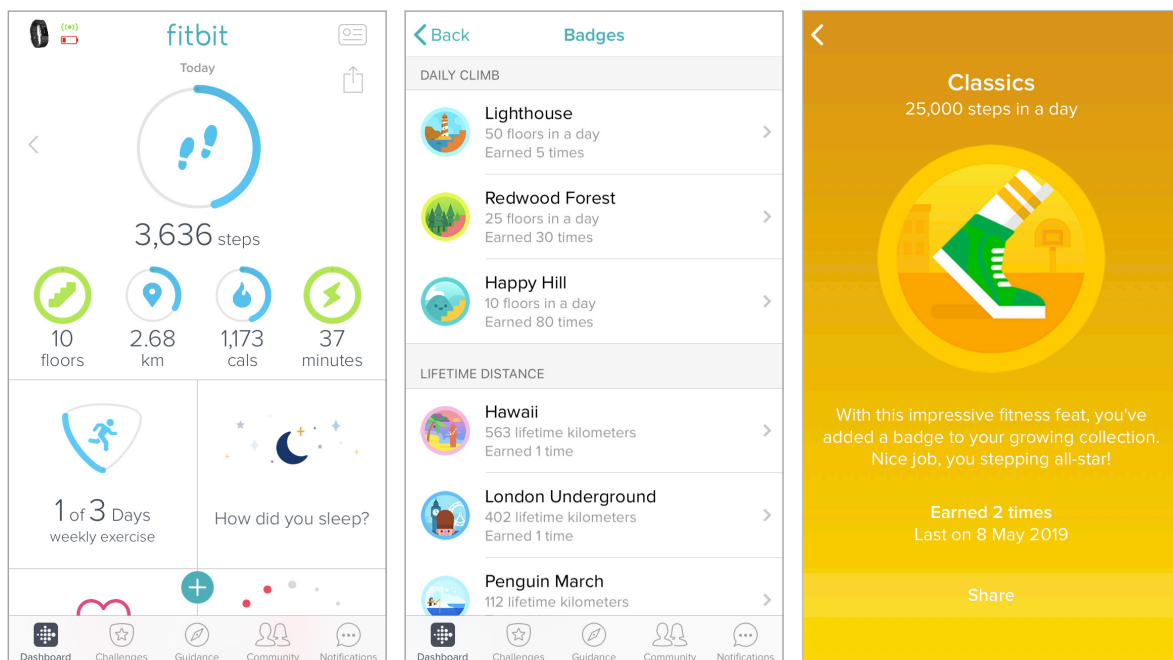


Figure 3. Screen images of the Fitbit app: Dashboard tab (left), a list of badges (right), and detail of a badge (right)

3.2.3. Garmin Connect app from Garmin

Garmin provides a mobile and web version of Garmin Connect, a tool for monitoring, analysing, and sharing health and fitness data. Garmin Connect is available to use once paired with a compatible Garmin activity trackers and smartwatches. While it supports more or less the same functionality as Fitbit does, its virtual reward system is more complex and multi-layered than that of Apple and Fitbit, in which different game design elements—badges, points, and levels—are combined into one achievement system named as recognition program. It offers a wide range of badges for different workout types, health-related activities, and even some tutorials. Both achieved and achievable badges are viewed in the mobile as well as web version of Garmin Connect. The most distinctive feature of Garmin's reward system is that it assigns different amounts of points into each badge based on the difficulty of the requirement. By accumulating the points, users are able to proceed to the next level. In comparison to that of Apple and Fitbit, Garmin's virtual reward system turns life into a game.

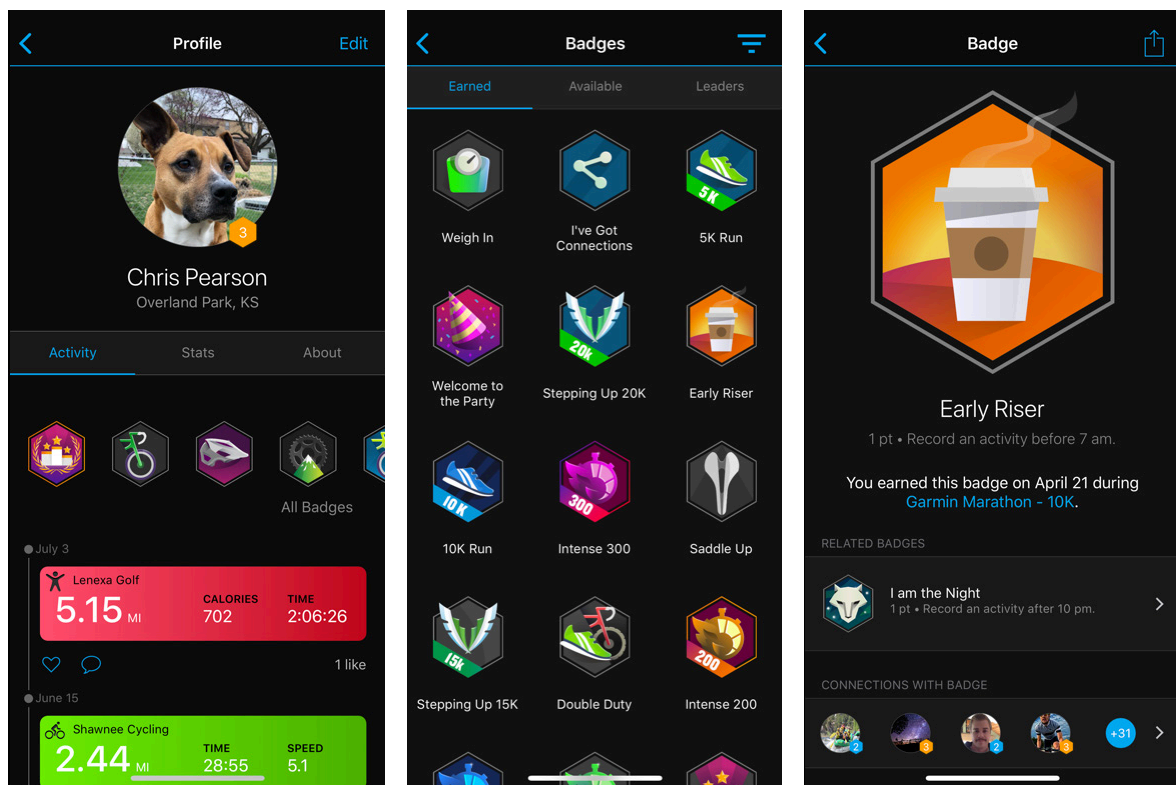


Figure 4. Screen images of the Garmin Connect app (Pearson, 2018): Profile (left), a list of badges (middle), and detail of a badge (right)

3.2.4. Samsung Health app from Samsung

Samsung Health app is all-in-one mobile platform where all health- and activity-related data from smartphones, activity trackers, and compatible devices are integrated, delivering the insights for a healthier lifestyle. It encourages users to determine three daily goals for the minutes of activity, intake of calories, and hours of sleep. It also incorporates a digital badge system to boost people's motivation and enjoyment, however, the underlying mechanism for the reward system is somewhat different from previous ones. While it shares some common traits with that of Apple Watch in that both focus on users' goal attainment, the biggest difference between the two is that Samsung Health allows a user to set a target goal before engaging in exercise, and then awards a badge when he or she reaches that goal. As with other platforms, users can see the collection of their rewards on their smartphones.

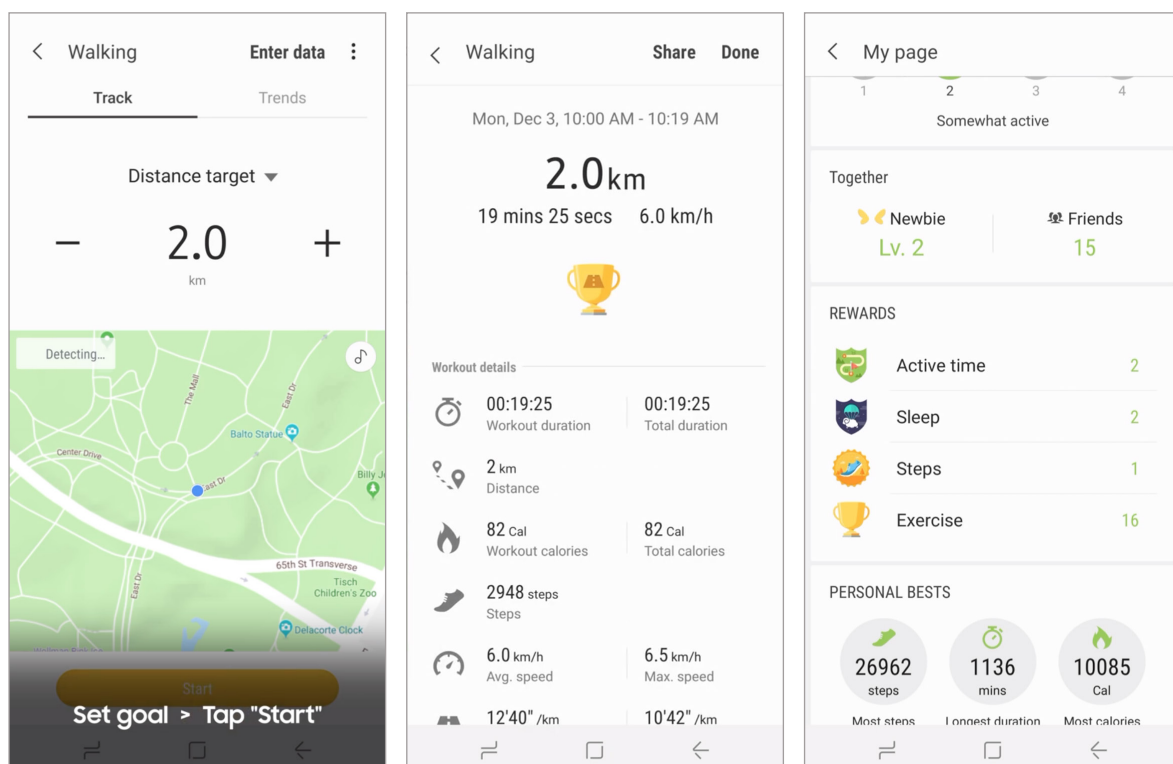


Figure 5. Screen images of the Samsung Health app (Samsung, 2019): exercise target setting (left), achievement of a badge (middle), and a list of badges (right)

3.2.5. Google Fit app from Google

As briefly introduced earlier, Google Fit app is a mobile platform for health and activity tracking provided by Google since 2014. With the recent redesign of the app, Google introduced two new metrics, Heart Points and Move Minutes, and a points system based on them, following the activity recommendations of the World Health Organization (WHO). According to the guidelines established by the organization (World Health Organization, 2010), adults should have 150-minute moderate exercise or 75-minute vigorous exercise at the minimum level. Heart Points and Move Minutes represent the intensity and the amount of physical activity, respectively. Heart Points give credit for activities that get the heart pumping harder such as brisk walking or jogging while Move Minutes are given for all of the activities that users engage in. In terms of Heart Points, the system awards points differently based on the intensity of exercise—one point for each minute of moderate exercise, and two points for each minute of vigorous exercise. Users are encouraged to set daily goals for both points, and the progress of goal attainment is represented by two concentric circles. While this simple and abstract reflection on daily activity level is available on both smartwatches and smartphones, the detailed information of each exercise is given only on the mobile version of the app.

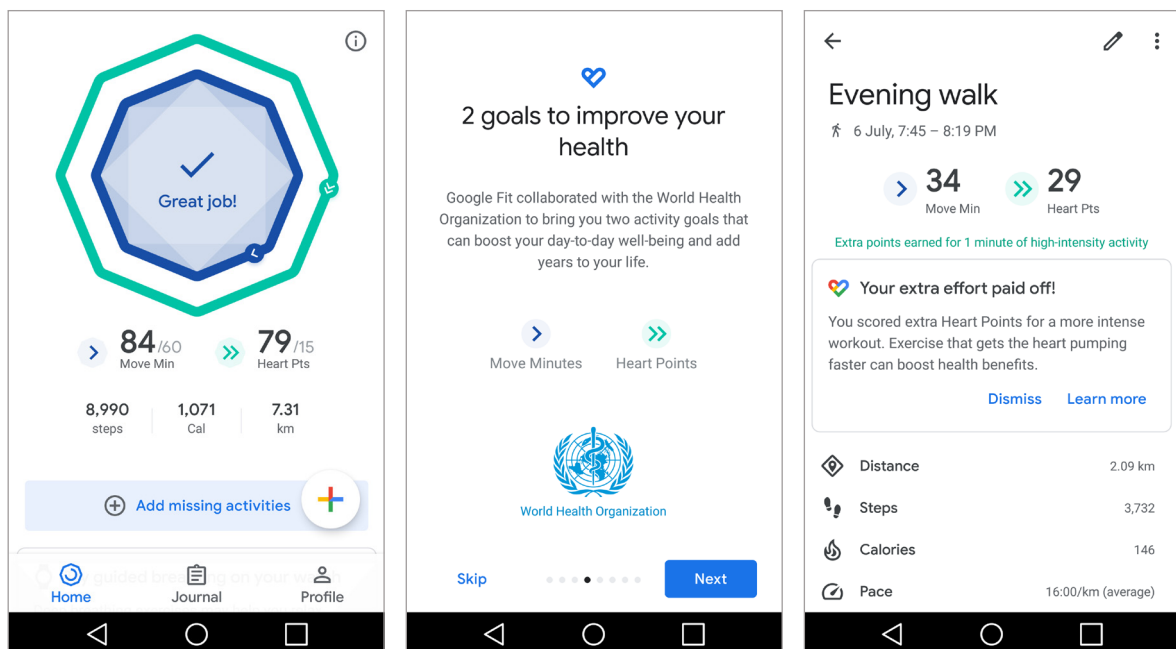


Figure 6. Screen images of new Google Fit app: Home (left), guidance on the points system (middle), and detail of an activity (right)

3.3. Comparative Analysis

In the previous section, I briefly reviewed the virtual reward systems employed in activity tracking systems of five major companies. Table 2 shows the summarized results of the review. During the review, the similarities and differences between the systems became apparent in terms of several points. First of all, Apple and Samsung focus more on users' goal attainment, whereas Fitbit and Garmin award badges based on the amount of measured data. Moreover, there are differences in the range of individual rewards between the systems. Apple and Garmin provide various types of rewards by applying them to diverse contexts while those of Fitbit and Samsung are somewhat limited to a few types of activity data. Also, the virtual reward system of each company has its own unique characteristics. Apple awards different digital badges according to the achievement rate of daily activity goal. On the other hand, Samsung allows users to set their target activity goal each time when users try to engage in physical activity. In addition, it only cares about whether to achieve the manually set goal, not counting the achievement rate as Apple does. Even if they do have the same focus, the content and logic of rewards can be significantly different. Fitbit and Google are trying to provide additional information with their virtual rewards. Fitbit uses distance-related, real-world information to metaphorically present how many miles users have walked since they started tracking. By earning those badges, users are actually able to learn something interesting. On the other hand, Google delivers more useful information than Fitbit. By employing the WHO's activity recommendations in its points system, users are enabled to be aware of whether their activity level meets the WHO's guideline, which could help managing health.

Table 2

Comparison between virtual reward systems in activity trackers of companies

	Apple/Activity app	Fitbit/Fitbit app	Garmin/ Garmin Connect app	Samsung/ Samsung Health app	Google Fit/Google Fit app
Type of Reward System	Badges	Badges	Badges / Points / Levels	Badges	Points
Concepts	Daily goal-based rewards	Measured data-based badge system	Measured data-based badge system	Individual activity target-based rewards	Health-related points system
Types of Individual Reward	<ul style="list-style-type: none"> • Daily goal attainment • Lifetime goal attainment • Best record (streak, new) • Workout • Challenge • Competition • Limited edition 	<ul style="list-style-type: none"> • Daily steps / floors • Lifetime distance / floors • Weight • Challenge 	<ul style="list-style-type: none"> • Daily steps • Lifetime distance • Best record (streak, new) • Workout • Active minutes • Challenge • Limited edition • Tutorials / system features 	<ul style="list-style-type: none"> • Daily goal attainment: steps / active minutes • Individual activity goal attainment 	<ul style="list-style-type: none"> • Intensity of exercise • Active Minutes
Unique Points	<ul style="list-style-type: none"> • Achievement rate of daily goal • Link between three (Move/Exercise/Stand) daily goal 	Metaphorical representation of lifetime badge	Digital badge system highly intertwined with points and levels	Manual goal-setting for individual activity	Different points based on the intensity of exercise

04

User Studies: Autoethnographic Diary Study

While differences in design approaches to virtual reward system became discernible through comparative analysis of the current designs of activity tracking systems, little is known about how these different characteristics affect users' real-life experiences with virtual rewards. Also, it was confirmed that there is still a need for further investigation of the influence and role of virtual rewards in physical activity practices since the findings from the reviewed literature indicated a mixed reception to them. To fill the knowledge gaps described above, I conducted an auto autoethnographic diary study and an online survey with existing users of activity tracking devices. In this and the following chapter, the detailed process and findings from both user studies are described.

4.1. Methods

4.1.1. Research Materials

In this study, I carried out a diary-based autoethnography with three different activity tracking devices: one activity tracker which is the Fitbit Charge 2 and two smartwatches which are the Apple Watch Series 4 and the Asus ZenWatch 2 (Figure 7). From the desk research on the current designs of virtual rewards in activity tracking systems, it became evident that current virtual reward systems are largely divided into two categories —badges and points—and that there are mainly two different logics to award badges to users—based on either the attainment of daily goals or predefined thresholds of measured data. Considering a company's market share in wearable devices and the results of comparative analysis of current designs, the Google Fit, Activity, and Fitbit apps are regarded as the most representative cases of a point-based reward system, goal-based badge system, and data-based badge system, respectively. Based on this consideration, three devices, Apple Watch Series 4, Fitbit Charge 2, and Asus ZenWatch 2, were selected since they were obtainable from the faculty or acquaintances. Each device is compatible with the Activity, Fitbit, and Google Fit apps, individually.



Figure 7. Images of research materials: Fitbit Charge 2 (left), Apple Watch 4 (middle), and Asus ZenWatch 2 (right)

4.1.2. Data Collection

Data was collected in two ways: by developing retrospective accounts (Duncan, 2004) for the experience of the Fitbit activity tracker, and by writing diary entries for that of the Apple and Asus smartwatches as reflections-in-action (Duncan, 2004). To be specific, the Fitbit charge 2 was used by the author in two separate periods: a one-week trial in May 2018 and two months of use from mid-March to mid-May in 2019. The device was rented from the Aalto Takeout, where students are able to borrow various equipment for study purposes, during both periods. The purpose of the one-week trial in 2018 was to have a first-hand experience

of activity trackers as part of a course. Other two teammates and I used the Fitbit charge 2 for a week focusing on how much quantified data correctly represents the quality of activities. In the trial period, we shared findings, thoughts, and feelings through a mobile messenger app. While the first trial was done for study purposes, the experience in the second period was intended for personal use. I had been using a Mi Band 3, an affordable activity tracker manufactured by Xiaomi, to monitor a regular exercise routine and sleep quality since the end of 2018. After a couple of months of use, I decided to try the Fitbit Charge 2 again with the expectation of more accurate measurement of data. Since the device had already been adopted and used before the plan of this research was constructed, the experience with virtual reward systems employed in the Fitbit app was written retrospectively (Figure 9). To bring a greater detail of experience, multiple sources of evidence were utilized such as a calendar, emails, records of push notifications, chat messages with teammates, and screens of the app. After that, a diary-based autoethnography with two other smartwatches, Apple Watch and Asus ZenWatch was conducted for two months in total, a month for each device, from early June to early August in 2019. Reflections-in-action were collected using the 'Notion' app, a cross platform note-taking tool available to use in both mobile and desktop environments. In the app, a template for a diary entry was manually produced and used to ensure the consistency of all data entered. Diary entries were not written on a regular basis but made when the experience related to virtual rewards was inspiring in either a positive or negative way, by which irregular intervals between the entries can be explained. To vividly capture the experiences, short descriptions were promptly noted in situ using the mobile version of the 'Notion' app with relevant screens on the watch and the smartphone captured (Figure 9). Details were added in the desktop environment later on that day. Each diary entry included descriptions of inspiring experiences, captured screens of the smartwatches and mobile phones, and comments on the researcher's thoughts and feelings. In some cases, ideas and suggestions were also added. While the Notion app was used in both cases as a software tool to collect diary entries, there was a difference between the two cases in terms of the device environment. Since the iPhone has been used in my daily life, I was able to easily add the Apple watch as a companion device and use the Activity app in a naturalistic setting. However, a workaround was needed in case of the Google Fit app. While a recent update of Google Fit app made it available on iOS, it was confirmed that some features were limited in the app for iOS devices. Therefore, to fully experience all the features of the Google Fit app, I carried another Android smartphone along with the iPhone at all times and connected it to the Asus ZenWatch 2.

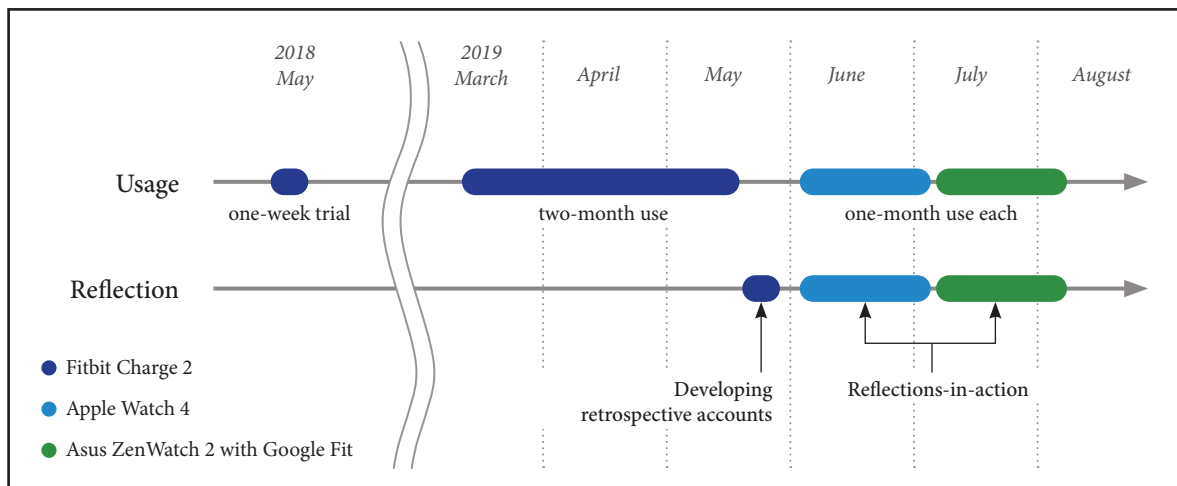


Figure 8. Timeline with an overview of autoethnographic research

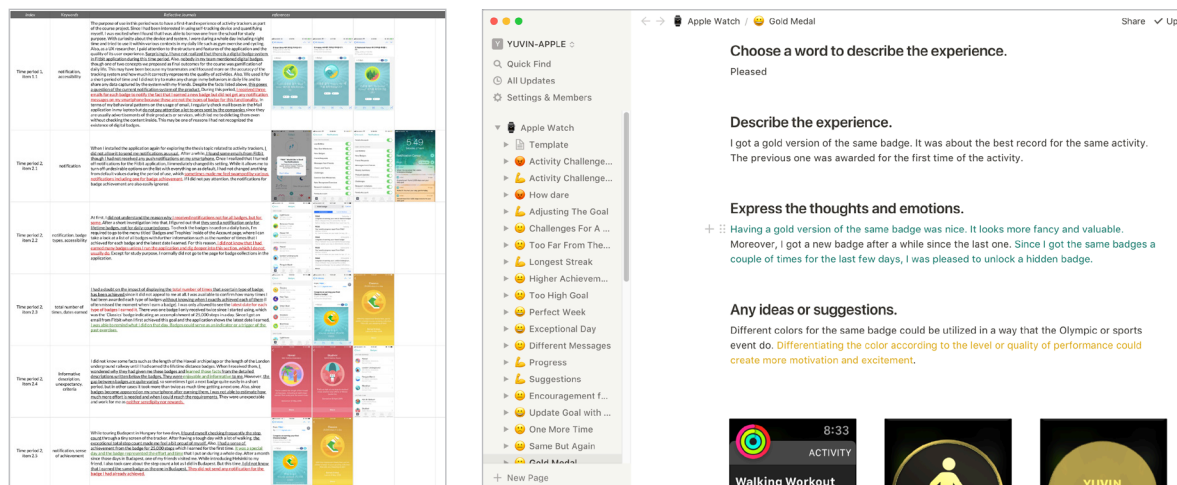


Figure 9. Collected data from autoethnography: retrospective accounts for Fitbit (left), reflections-in-action through the Notion app for Apple Watch and Google Fit (right)

4.1.3. Data Analysis

As a result, 12 retrospective accounts about Fitbit's virtual reward system were gathered, among which 11 items came out of the experiences in the second period. Also, a total of 49 reflections-in-action were collected from the use of the Apple Watch and Asus ZenWatch while 33 diary entries were made for the Activity app and 16 entries were for the Google Fit app. Both retrospective accounts and reflections-in-action were thoroughly reviewed, and the statements which described positive and negative influences of a virtual reward system on my behavior, attitude, and emotion were extracted. In addition to 79 positive statements and 54 negative ones, 21 ideas or suggestions were also prepared for further analysis. Therefore, 154 items in total were analysed in more depth by affinity diagramming (Lucero, 2015).

4.2. Findings

Through affinity diagramming, 154 statements were thematically grouped into 29 categories and these were further developed into 6 themes (Figure 10). The themes identified include 1) *content*, 2) *aesthetic*, 3) *access*, 4) *trust*, 5) *communication*, and 6) *goal-setting*. The first four themes cover different aspects of experience with virtual reward systems while the other two are the relevant features that interact with a reward system for sustained user engagement. In this section, I present the primary findings from my experiences with three different virtual reward systems in current activity trackers along six themes above.

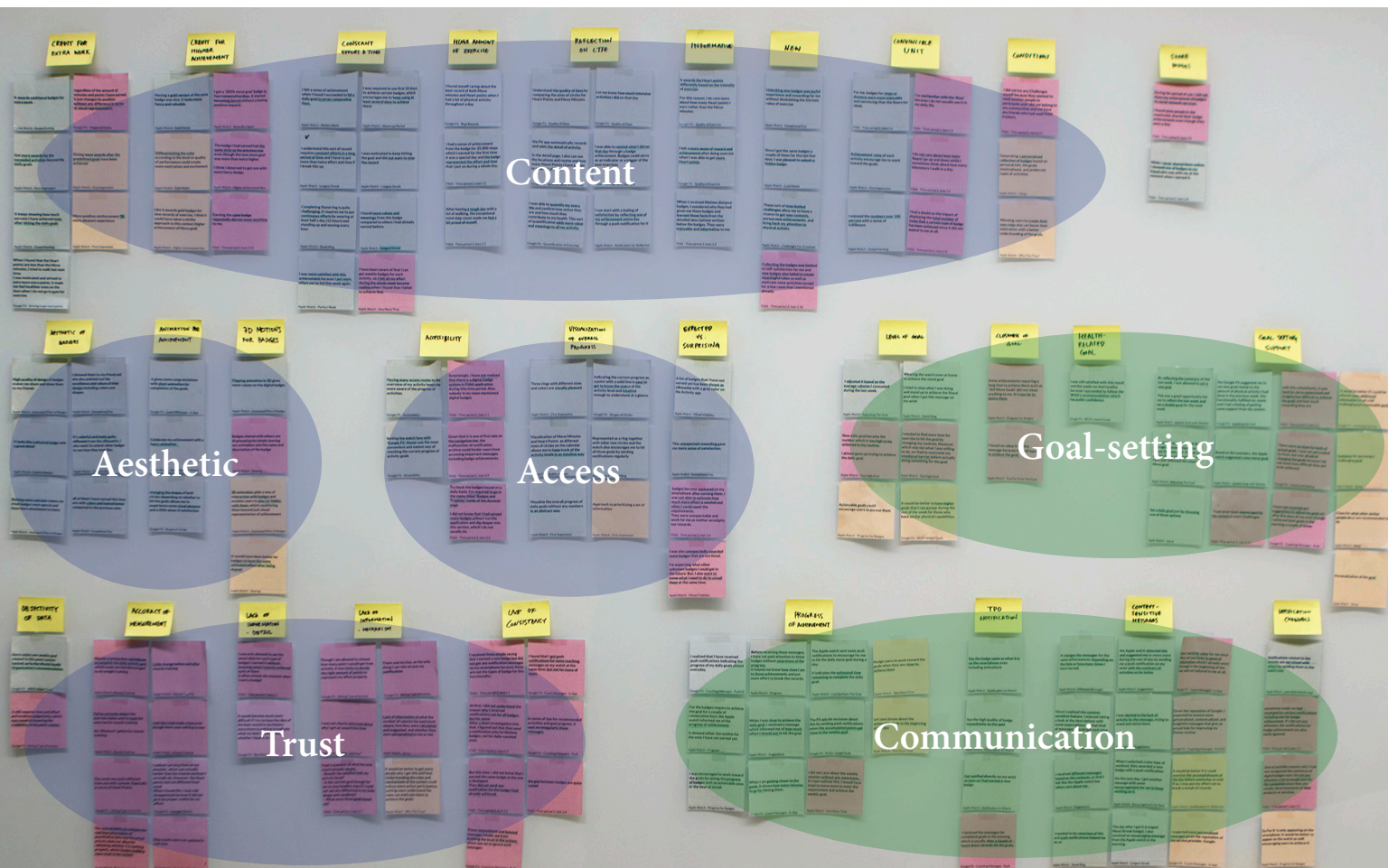
4.2.1. Content

As pointed out from the prior study (Antin & Churchill, 2011), not all badges or points were valuable and motivational to me. However, the types or characteristics of virtual reward that were highly appreciated and meaningful have become clear during the study.

Proper credit for effort and time

First of all, the proper credit for effort and time was important. After a while using Google Fit in my everyday routines, I found myself caring much more about Heart Points than Move

Figure 10. Result of affinity diagramming with 6 major themes



Minutes. Unlike Move Minutes, Heart Points awards different points per minute depending on the intensity of exercise I engage in. In other words, the more effort I put into exercise, the more points I could get. This clear connection between effort and outcomes made me feel like I was properly acknowledged by the system and keep motivated even after I reached daily goals. Another episode in the experience with the Apple Watch also shows the importance of the appropriate acknowledgement, but in an opposite way:

Reflections-in-action for Apple Watch, Date: 17/06/2019

I received the New Move Goal badge once I hit the new move goal which was set higher. However, the badge I had earned had the same style as the previous one with only some information updated. Given that the new move goal was more than twice higher than the previous goal, I think I deserve to get a badge with a fancier design. Since the Apple watch awards gold badges for new records of workout, I think a similar approach could have been taken in this context as well.

Information of the impact on health

Moreover, I appreciated the informational value of virtual rewards. One of the major merits of adopting activity trackers into our daily lives is to have better awareness of ourselves and activities that we engage in. However, the novelty of monitored information has worn off over time and many people fail to gain insights from the tracked data (Jarrahi et al., 2018). To lessen the burden of the interpretation of data, a composite metric can be provided (Zichermann & Cunningham, 2011). In this sense, Heart Point is also considered as this type of metric. While the NikeFuel point was a more arbitrary metric, Heart Point is more sensible and informative to ordinary people since it has a direct link to the activity recommendations from the WHO (World Health Organization, 2010). This enabled me to have a clear connection between the points I got through exercise and their impact on my health. In other words, the amount of Heart Points informed how much influence a certain type of physical activity has on my health improvement. This is well described in the following diary entry:

Reflections-in-action for Google Fit, Date: 07/07/2019

When I went shopping, I looked around the store for an hour. However, I received only a few Heart points, which was actually fair enough considering what I did in the store. I walked slowly and stopped to see products many times, which was lower intensity of exercise compared to the walking activities in another day. For this reason, I do care more about how many Heart points I earn rather than the Move minutes, and I felt a greater sense of reward and achievement after doing exercise when I was able to get more Heart points.

Challenging

It is common that we find greater value in the compensation for competing more difficult, challenging tasks. What does it mean by challenging in the context of physical activity tracking? I found it challenging when I had to make a constant effort to keep hitting the daily activity goals, break the best record, and do a considerable amount of exercise in a day. The virtual rewards I earned for these challenging tasks were more valuable, motivating me to keep engaging in physical activity with a sense of achievement and satisfaction:

Reflections-in-action for Apple Watch, Date: 17/06/2019

After hitting the daily move goal for eight consecutive days, I had finally earned the Longest Move Streak badge. I found more values and meanings from this badge compared to others I had already earned before. Also, I was motivated to keep hitting the goal and did not want to stop the record. This was because I understand this sort of record requires constant efforts in a long period of time and I have to put more than twice the effort and time if I break this up.

In addition, the experiences of receiving rewards for huge achievements were so memorable that I was able to vividly recall the memories of those times through the badge or the records of points:

Retrospective accounts for Fitbit, Time period 2, item 2.3

I had a doubt on the impact of displaying the total number of times that a certain type of badge has been achieved. There was one badge I only received twice since I started using Fitbit, which was the 'Classics' badge indicating an accomplishment of 25,000 steps in a day. Since I got an email from Fitbit when I first achieved this goal and the application shows the latest date I earned, I was able to remind what I did on those two days. Badges could serve as a trigger of the past exercises.

However, since I found greater value in the badges awarded for consistency, I was also discouraged when I failed to achieve that due to the circumstances beyond my control or lack of awareness of progression:

Reflections-in-action for Apple Watch, Date: 10/06/2019

I failed to earn the 'Perfect Week' badge due to yesterday. I did not care about this without any awareness of weekly mission. I was disappointed with this because I'm pretty sure I must have moved more had I known it earlier. I have been aware of that I can get weekly badges for each activity, so I felt all my effort during the whole week became useless when I found that I failed to achieve that.

New

In the course of using activity trackers, I have earned the same badges many times, especially the ones that require a relatively low level of commitment. Once I started using each device, I was able to unlock new badges often. However, the chances to achieve new badges decreased over time while I had repeatedly received the same badges by hitting the same goals. The novelty effects of these badges wore off at this point, and they failed to yield any positive influences on my behavior and emotion. While both Fitbit and Apple indicate the total number of times I earned each reward, that number also did not mean anything to me. In this context, Apple's time-limited challenges did serve as a good refresher:

Reflections-in-action for Apple Watch, Date: 20/06/2019

I received push notifications to inform of some time-limited challenges. Some of them are ones for celebrating the annual events which last for only one day. In other cases, there are monthly challenges which have requirements for activities in a month. These sort of time-limited challenges allow me to have a chance to get new contents, pursue new achievements, and bring back my attention to physical activity.

Aligned with my interest and context

As stated in the reviewed literature about gamification (Deterding, 2011; Hamari et al., 2014; Nicholson, 2012; Vassileva, 2012), the effects of my interest and context on the outcomes were also confirmed during my study. Since I do not care about the flights I climbed in my daily life and I have some doubts about the accuracy of measurement of floors climbed, I paid little attention to the rewards for that as well. In terms of step counts, I only care about it when I have an exceptional amount of physical activity, such as during an excursion or travel. While I'm not usually interested in the measured step counts or floors, I was encouraged by the achievement rate of daily goals on the Apple Watch. So, the rewards awarded by the Apple Watch were mostly more meaningful to me than the ones by Fitbit. In short, the value of virtual rewards was proportional to the degree of my interest. Also, the environment where I engage in physical activity tracking determined the types of rewards I was attracted to. During the study, there was no one around me who had regularly used an activity tracking device in the real world, and I did not belong to any online community. Affected by this context where I was situated, I did not try any of Fitbit's challenges since I needed another user to take part in most of challenges. This was also true for the Apple's competitions as well.

4.2.2. Aesthetic

In comparison with physical forms of reward, digital rewards have their own pros and cons. Though they cannot bring tangible benefits to users' real life, digital rewards can be more

dynamic and promptly respond to a user's actions (Wang & Sun, 2011). Also, they can be graphically materialized as realistic as their counterparts. How to utilize these attributes could affect users' playfulness and perception on virtual rewards.

Skeuomorphs vs. Flat

There are some differences in the current designs of digital badges between Apple and Fitbit. One significant difference is that the badges of Fitbit are flat illustrations, whereas those of Apple display skeuomorphic metal textures, looking like real-world badges. Skeuomorph is defined as "an element of a graphical user interface which mimics a physical object" ("Skeuomorph", 2019). Given that a flat design is a common approach to designing digital badges taken by many companies such as Fitbit, Garmin, and Samsung, a skeuomorphic design in this context is distinct from others. When I earned my first badge with the Apple Watch, I experienced how this different approach to design could influence my behavior and attitude towards digital badges:

Reflections-in-action for Apple Watch, Date: 03/06/2019

I was able to see the backside of a badge where my name and the date I earned were etched. Etching name and date makes me treat badges more special and have more attachment to them. Since it is made with high quality of design, I wanted to share and show badges to my friends.

Reflections-in-action for Apple Watch, Date: 04/06/2019

The first couple of badges were designed with silver color. It did not appeal that much to me but the 'Move Goal 200%' badge was different. It's colorful and looks quite different from the silhouette which I was able to see before earning it. It looks like a physical badge with a great detail. I also want to unlock other badges to see how they look like.

I should admit that this sort of skeuomorphic design will not be universally appreciated. The preferences for graphic styles are different and varied from person to person. However, it implies that the high quality of graphical representation could add additional value to badges and appeal to users' emotions.

Interactivity

Another notable difference between the badges of two companies is in terms of interactivity. While Fitbit's badges are flat and static, those of Apple are more spatial and dynamic. In other words, while people are only allowed to appreciate the graphics of badges with Fitbit, the users of Apple Watch are enabled to flip and fiddle with badges on their watches as well as smartphones, as we normally do with physical objects. This sort of interactivity might

be trivial but can make digital badges more playful by bringing them beyond just a visual representation of achievement.

4.2.3. Access

Easier access, more awareness

While different types of devices allow different levels of accessibility to the virtual reward systems, easy access to the reward system can lead to more immersive user experience with digital rewards. While Google Fit and Apple Watch enable users to check their rewards—points and badges, respectively—on their wrist by offering a watch face theme or a menu in the watch version of app, Fitbit allows access to the collection of badges only from the smartphone app. Also, since it informs users of new achievements of some badges via email and push notifications on smartphones, there was no way to access the rewards on the activity tracker during my experience of Fitbit Charge 2. These made a huge difference in my level of awareness about virtual rewards, which was illustrated through the following episodes:

Retrospective accounts for Fitbit, Time period 1, item 1.1

Surprisingly, I have not realized that there is a digital badge system in Fitbit application during this time period even though I received three emails regarding new achievements of badges.

Retrospective accounts for Fitbit, Time period 2, item 2.2

To check the badges issued on a daily basis, I'm required to run the application and go to the menu titled 'Badges and Trophies' inside of the Account page, which I do not usually do. For this reason, I did not know that I had earned so many badges.

4.2.4. Trust

Reward is a thing given in recognition of service, effort, or achievement ("Reward", 2019). In order for rewards to be properly accepted, the trust between receivers and givers needs to be built. During the study, three factors that affect my trust in the reward system were identified.

Objectivity

In activity tracking systems, virtual rewards are awarded by the service providers based on the activity data. Due to the limited capability of the current wearable technologies, not all activities are accurately tracked by the devices. As a workaround, some companies allow users to manually create and edit the records of their physical activity. While improper credits from

tracking systems can discourage engagements with physical activity, manual input needs to be approached with caution since it may undermine the objectivity of the system, thereby lowering the value of rewards. It was depicted in the following diary entry:

Reflections-in-action for Google Fit, Date: 10/07/2019

One day, I forgot to choose the type of exercise in advance, relying on the auto-detection functionality. The result was quite different from one with a preset, Especially in terms of Heart Points. Though I am allowed to choose how many points I would get from an activity, it was tricky to decide the right amount of points to represent my effort properly. There was no clue, so the only thing I can rely on was my justification. I could make an assumption of the value with reference to the previous results, but it requires time and effort and involves subjectivity, which may result in lowering the credibility of the point system.

To avoid this, the companies that have adopted a digital badge system clearly note that badges are not granted for manually logged activities (Fitbit Inc., 2019; Garmin Ltd., n.d.)

Consistency and Transparency

Users feel their experiences are consistent when the mechanism of a system is constantly recognized in their use. When the outcomes of their actions are different from what they expect to get, users get frustrated and may lose their confidence in the system. During the study, I experienced the lack of consistency of reward systems at some points. Some were caused by the lack of detailed information while some were brought about by the incomprehensible mechanism. This can be improved by making the system more transparent to its users with the provision of sufficient information of the rules and mechanisms of a system, which could consequently induce more active participation:

Retrospective accounts for Fitbit, Time period 2, item 2.2

At first, I did not understand the reason why I received notifications not for all badges, but for some. After a short investigation into that, I figured out that they send a notification only for lifetime badges, not for daily-counted ones.

Reflections-in-action for Apple Watch, Date: 07/06/2019

I recorded an outdoor walking a couple of times, but I have not received this reward until today. I was glad to get a new badge but wondering why this time. It may be because previous workout was in warm-up period or because of the duration of exercise. In anyways, I was not clearly informed about it. It would be better to get more details of why and how I get a badge.

4.2.5. Communication

All activity trackers I experienced have been using various channels to communicate with their users for many purposes, and those channels are used for virtual reward systems as well. However, differences exist between the systems in terms of types of communication channel, content in the messages, and the timing of a message.

Notification channel

All three tracking systems have used different communication channels. While the Activity app uses a single channel to communicate with its users regarding virtual rewards, the Google Fit and Fitbit app use multi-channels: the Google Fit app displays messages on both the smartwatch and the smartphone while the Fitbit app sends messages via email in addition to a push notification on the smartphone. However, given that there are many other applications on smartphones and the activity data is mainly tracked through wearable devices, a watch-only notification can be more effective to attract users' attention and stay focused on the activities they engage in. An adverse effect of multi-channel notifications was found in one of retrospective accounts:

Retrospective accounts for Fitbit, Time period 2, item 2.1

While the app allows me to turn off undesirable options on the list of possible notification messages with everything on as default, I had not changed anything from default values during the period of use. As a result, I sometimes felt swamped by notifications including those for the badge achievement. Without careful attention, the notifications for badge achievement were also easily ignored.

Progress

I was notified of the progress for an achievable badge and the attainment of daily and weekly goals on my watches or smartphones while using the Apple Watch and Google Fit. Considering some rewards call for users' continued focus and constant efforts and there exist different levels of thresholds for daily activity, this sort of messages can help stay motivated to achieve a challenging, but achievable goal of reward, which consequently results in increased physical activity:

Reflections-in-action for Apple Watch, Date: 26/06, 13/06/2019

The Apple watch informed me of the progress for the badges that require a streak of goal achievement. In addition to the text messages, it displayed the grey outline of the badges I am aiming for. Before receiving these messages, I have not paid attention to

these badges without awareness of the progress. It helped me know how close I am to those achievements and put more effort to keep up with the records.

Time, Place, and Occasion (TPO)

The impacts of virtual rewards can vary greatly depending on when and in what context the messages for new achievements are delivered. It can be more influential and give a greater sense of achievement and pleasure when users are enabled to appreciate new achievements on the spot in real time. In other words, the novelty effect of rewards may wear off if they are delayed or delivered in other contexts. The importance of the timing of a message was well described in the following diary entry:

Reflections-in-action for Apple Watch, Date: 03/06/2019

I received a push notification for the badge that I just earned with the Apple watch, and I was able to even see them. It was great to get notified directly on my wrist as soon as I had earned a new badge, which I did not expect to get. Also, I was enabled to see the badge in the same way as I do on the smartphone even including animations. Seeing a high-quality badge immediately on the spot was an enjoyable experience and I felt a sense of satisfaction.

4.2.6. Goal-setting

Challenging, but achievable goal

While the badge system of Fitbit has nothing to do with daily activity goals, that of Apple Watch is closely intertwined with them, especially the daily Move goal. Since the Activity app awards different badges to users based on the achievement rate of the Move goal, the lower the daily Move goal is, the easier users are able to earn badges. In this sense, rewards are strongly linked to the goal-setting. The Apple Watch supports to adjust the daily goal every week with a weekly summary of physical activity. In addition, it suggests a new Move goal based on the summarized data. When I first adopted the Apple Watch, I set a low goal following the default value, so that it was quite easy to hit the goal and earn badges for the first couple of weeks. However, it had been completely changed with the increased Move goal as described in the snippet of the entry:

Reflections-in-action for Apple Watch, Date: 17/06/2019

I was suggested to set a daily goal to 720 calories by the Apple watch, which is twice higher than the previous goal. The problem was I backed into a routine after a couple of exceptional days, which led to lowering the activity level. New daily goal became the

number which is too high to be achieved in my routine. I almost gave up trying to achieve the daily goal. I needed to find more time for exercise to hit the goal by changing my routines, however, which was not what I was willing to do.

Setting an appropriate level of activity goal is quite challenging in real-life contexts. As with the Apple Watch, if virtual rewards have a close connection with a daily activity goal, the quality of experience with virtual rewards can be influenced by the goal-setting practices.

Health-related milestones

Unlike digital badges, Move Minutes and Heart Points in the Google Fit app are continuously accumulated from the start of physical activity. Although there is no threshold for rewards in point-based reward systems, Google Fit applies the activity recommendations from the WHO (World Health Organization, 2010) to its weekly activity goal for all users. When users get close to the WHO goal, Google Fit informs users of this and celebrates the completion of the goal once users earn 150 Heart Points. This kind of health-related milestones can add practical value to virtual points as indicated in the entry below:

Reflections-in-action for Google Fit, Date: 09/07/2019

I hit this goal in the first two days of the week, which was quite easy to complete. But I was still satisfied with this result and this made me feel healthy because I succeeded to follow the WHO's recommendation which has public confidence. It would be better to have higher goals that I can pursue during the rest of the week for those who have similar physical capabilities.

4.4. Summary of Findings

From 4-month length of autoethnographic diary study with three activity tracking systems, the detailed interactions with virtual rewards in real-life practice were illustrated. The results identified 15 opportunities under 6 themes—1) *content*, 2) *aesthetic*, 3) *access*, 4) *trust*, 5) *communication*, and 6) *goal-setting*—that could contribute to improving the quality of virtual reward experience. In detail, virtual rewards are considered more valuable if they 1) give proper credits for effort and time, 2) inform the impact on health, 3) are challenging, 4) new, and 5) aligned with personal interest. Also, the playfulness of virtual rewards can be differed by the way of graphical representation and interactivity. For more reliable and convenient use of virtual reward systems, accessibility, transparency, and consistency of the system need to be taken into account. In addition, the effectiveness of virtual rewards can be enhanced through appropriate communication. In my autoethnographic research, the effect was greater when information on progress is provided and rewards are delivered in real time on the spot. Lastly, the level of engagement with virtual reward systems can be increased by helping users set goals within the reach if virtual rewards are tied to daily activity goals and providing health-related milestones.

User Studies: Online Survey

My autoethnographic diary study with three activity tracking systems provided insights into what elements and properties construct the experience of virtual rewards, how different designs of reward system influence users' perceived values of virtual reward and their engagement in physical activity, and how virtual rewards interact with other features. Based on the in-depth understanding of current virtual reward systems gained through the diary-based ethnographic research, I conducted an online survey, seeking to inquire into how virtual rewards have influenced people's physical activity practices in the real world and how they are perceived and valued by their users.

5.1. Methods

5.1.1. Participants

Participants were recruited through Amazon Mechanical Turk (MTurk), which is one of the crowdsourcing platforms where people are contracted on a case-by-case basis and get paid for performing tasks remotely. The target audiences were those who have been using one of the following devices or apps: Apple Watch, Fitbit, and Google Fit. People who fit the criteria above were allowed to participate in the study and they were compensated 1.50 USD for completing the approximately 15-minute survey. To ensure the quality of responses, additional qualifications were also specified: a task approval rate at or above 95 percent and the completion of at least 500 approved tasks. In total, 336 individuals participated in the survey over three days. Among them, 216 participants out of them were qualified through screening questions and completed the full questionnaire. Out of those, 80 responses were rejected due to the suspicion about their integrity, such as giving duplicated answers to several open-ended questions or containing multiple answers that did not match the chosen tracking system. Although the remaining 136 responses were approved and paid, 23 of them were excluded due to the ambiguity of answers. Thereby, a total of 113 complete responses were further analysed.

5.1.2. Questionnaire

A questionnaire was created on Webropol, an online analysis and survey tool. In the survey, participants were asked about their personal experiences and opinions about virtual rewards—badges or points—offered by one of three activity tracking devices and their dedicated mobile apps: Apple Watch and Activity app, Fitbit smartwatches/trackers and Fitbit app, and Wear OS—formerly Android Wear—smartwatches and new Google Fit app. To reach target audiences, two screening questions were placed at the beginning of the survey, asking whether to use one of three activity tracking systems at present. Qualified participants were allowed to take part in the study while those who did not meet the criteria were terminated from the survey. The questionnaire consisted of three parts: *demographics*, *fitness tracking practices*, and *experiences with virtual rewards*. To understand who have completed the survey, I first asked basic personal information, such as gender, age, country of residence, and nationality, as well as the current activity levels of participants. It was followed by the questions about fitness tracking practices: duration of use, frequency of updating the daily goal of activity tracking system, types of data initially and currently monitored with an activity tracker, and prior and current motivation for activity tracking. After the inquiry into the qualities of participants, the questions focused on how they have perceived and experienced the current virtual reward systems in everyday practice. To get deeper insights into users' lived

experiences with virtual rewards, I asked participants several open-ended questions about a memorable experience with virtual rewards, names of valuable and unimportant badges or points, an experience of sharing achieved rewards and receiving real-world rewards, and suggestions for the improvement of virtual rewards in activity tracking systems. In addition to these free response questions, I also asked the level of familiarity and satisfaction with current designs of virtual rewards using five-point Likert scale. Moreover, in order to assess the perceived value and effect of current virtual reward systems, I asked participants to indicate their level of agreement to 24 statements—22 statements for Google Fit users—using seven-point Likert scale ranged from “strongly disagree” to “strongly agree”. The Likert scale items were constructed based on the evidence confirmed by the reviewed literature, the findings from the autoethnography, and GAMEFULQUEST, an instrument developed to measure a user’s gameful experience in systems and services (Högberg, Hamari, & Wästlund, 2019). While each item was intended to evaluate different aspects of user experience, they were thematically grouped into three: 7 items—5 items for Google Fit users—as to the quality of system, 7 items as to perception and attitude towards virtual rewards, and 10 items as to behavioral and psychological effects of virtual rewards (Table 3).

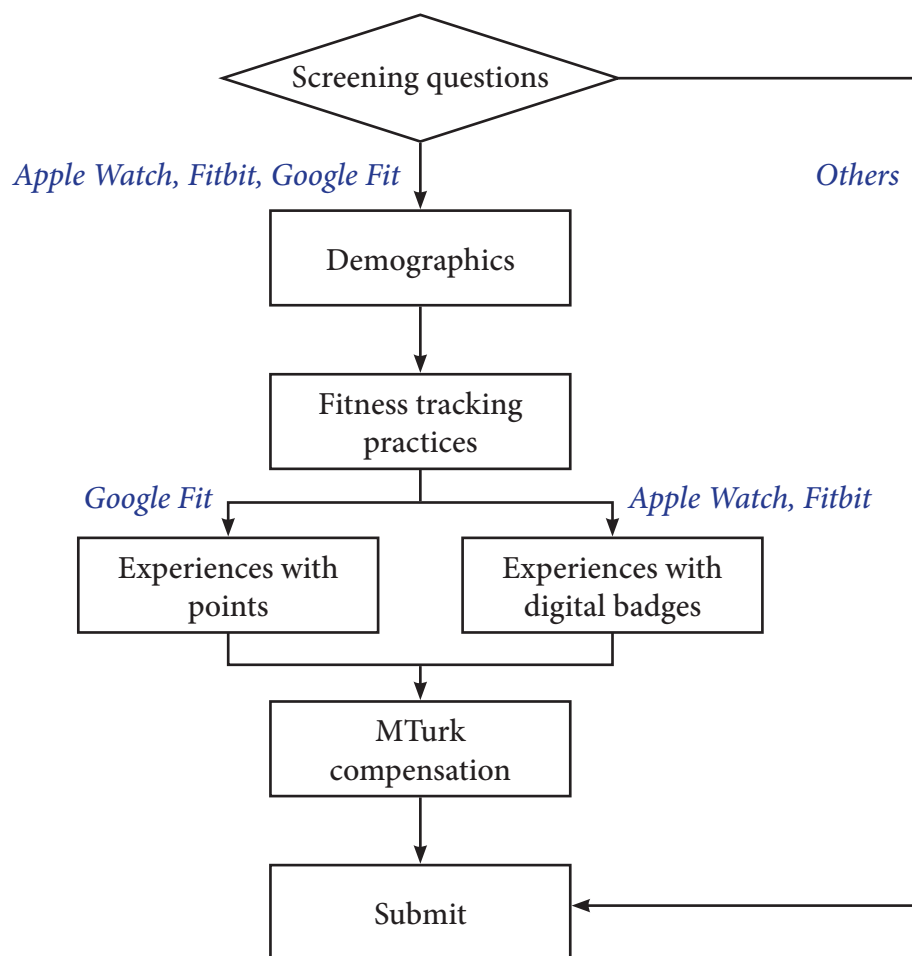


Figure 11. Overall flow of the survey

Table 3

A list of Likert scale items

Groups	Abbreviations	Statements	Inspired by
Quality of System	S_Aesthetic	Badges are aesthetically appealing	Autoethnography
	S_Access	It is easy to get access to the badges that I have earned	Autoethnography
	S_Ease	It is easy to understand how the reward system works and what I need to do	Autoethnography
	S_Trust	Rewards seem professional and trustworthy	Autoethnography
	S_Support	The system helps me to achieve the rewards with appropriate support	Autoethnography
	S_Frequency	I wish to get rewards more frequently	Autoethnography
	S_Diversity	I am satisfied with the current diversity of badges	Autoethnography
Perception and attitude	P_Emotion	I am glad when I earn a badge/point	Autoethnography, Fritz et al. (2014), Karapanos et al. (2016)
	P_Value	I find badges/points valuable	Autoethnography
	P_Credit	I think the badges/points represent my effort well	Autoethnography, Fritz et al. (2014), Harrison et al. (2015)
	P_Enjoy	Earning and collecting badges are fun and enjoyable	Autoethnography. Karapanos et al. (2016)
	P_Difficult	I am confident in my abilities to achieve them	Autoethnography, Zichermann & Cunningham (2011)
	P_Curiosity	Badges appeal to my curiosity	Autoethnography, GAMEFULQUEST
	P_Willingness	I'd like to continue to earn and collect badges	GAMEFULQUEST
Behavioral and psychological effects	E_Goal	<i>(Earning and collecting badges...)</i> Makes me feel like I have clear goals	Autoethnography, GAMEFULQUEST, Fritz et al. (2014)
	E_Progress	Confirms that I am making good progress	Autoethnography, Jarrahi, Gafinowitz, & Shin (2018)
	E_Motive	Helps me keep motivated to engage in physical activity	Autoethnography, Vooris et al. (2019)
	E_Pride	Makes me feel proud of myself	Karapanos et al. (2016)
	E_Feedback	Gives me useful feedback so I can adapt	Autoethnography, GAMEFULQUEST, Jarrahi, Gafinowitz, & Shin (2018)
	E_Enjoy	Helps me enjoy physical activity more	Autoethnography, Fritz et al. (2014), Vooris et al. (2019)
	E_Effort	Encourages me to put an extra effort to acquire them	Autoethnography, Fritz et al. (2014), Vooris et al. (2019)
	E_Social	Makes me feel like I am socially involved	GAMEFULQUEST
	E_Competetion	Makes me strive to be the best	GAMEFULQUEST
	E_Negative	Makes me feel bad about myself if I haven't achieved them in a while	Fritz et al. (2014)

5.1.3. Data Analysis

I adopted a mixed-method approach to analyse the data with different tools since the collected responses included both quantitative and qualitative data. For the responses to open-ended questions, a content analysis (Columbia University, n.d.) was performed using ATLAS.ti (Version 8.2.3). Long-form text data were coded and grouped together based on the codes. In some cases, a single open-ended text response contained multiple codes when a user mentioned two or more answers to one question. On the other hand, the responses to the closed-ended questions and Likert scale items were analysed quantitatively with SPSS (IBM SPSS Statistics, Version 25). The frequency of responses and percentages were calculated to identify evident trends in a data set. Furthermore, the differences and correlations between systems and user groups were also estimated with inferential statistics.

5.2. Results

In this section, I present the results of the survey. I start by characterizing people who participated in this study with demographic data and information about their use of activity tracking devices. Subsequently I illustrate how people perceive and interact with virtual rewards offered by activity tracking systems in their physical activity practices.

5.2.1. Participants Characterization

Of the remaining 113 respondents, 69 (61.1%) were male and 44 (38.9%) were female. Given that the gender ratio of wearable users in the US is close to fifty-fifty (EMarketer, 2019), that of the participants was slightly skewed, but still an acceptable division. The majority of respondents were from the US (96), followed by India (6) and other countries (11). Also, most of them lived in the US (102), and 6 respondents in India. In terms of age, more than half of participants (54.0%) were between the ages of 25 and 34, followed by 24.8% between 35 and 44, 15.0% between 18 and 24, and 6.2% between 45 and 64. Compared with the results of the recent report on the adoption of wearable devices (EMarketer, 2018), this study showed the same trends in age, but a higher proportion of the 25 to 34 age group was observed. This can be explained by the fact that people on MTurk tend to be younger and 60% of them are born after 1980 (Difallah, Filatova, & Ipeirotis, 2018). Regarding the current activity level of participants, approximately four fifths of them were frequently engaging in physical activity in their everyday life—68 (60.2%) were 3 to 5 times a week and 25 (22.1%) were 6 to 7 times. 19 participants (16.8%) reported doing physical activity once or twice a week, and only one participant (0.9%) responded that he or she rarely does physical exercise.

5.2.2. Fitness Tracking Practices

Usage of activity tracking device and app

When it comes to the currently using activity tracking device and service, 67 participants (59.3%) reported using a Fitbit smartwatch/tracker and the Fitbit app while 32 (28.3%) reported using an Apple Watch and the Activity app and 14 (12.4%) using a Wear OS smartwatch and new Google Fit app. This imbalance between these activity tracking systems can be explained by several reasons. First of all, the market share of Apple Watch and Fitbit is much higher than Wear OS compatible smartwatches. Second, the price range of Fitbit smartwatches/trackers (\$100-230) is relatively lower than that of Apple Watch (from \$400), so they can appeal to a broader public. In terms of the length of time since they adopted the current trackers, they were relatively evenly distributed. 25 respondents (22.1%) owned their trackers for 3 months or less, 33 (29.2%) for 4 to 6 months, 25 (22.1%) for 7 to 12 months,

18 (15.9%) for 1 to 2 years, and 12 (10.6%) for 2 years or more. Given that the use of activity trackers was abandoned within six months by one-third of users in the US (Endeavour Partners, 2014), both long-term (55)—7 months or more— and short-term users (58)—6 months or less— were almost equally recruited.

Motivations for adopting activity tracking tool

Some pre-existing motivations for the adoption of activity trackers were revealed through the content analysis of text responses. 31 participants aimed to improve health by increasing their activity level or losing weight while 16 desired to maintain health. 29 adopted for tracking their daily activity including step counts, and 22 for tracking exercise, and 8 for tracking biometric data. These 106 participants (93.8%) can fall into the purposive group, according to Karapanos et al. (2016). Only 7 (6.2%) adopted for other purposes such as using smartwatches or socializing with friends and family, belonging to the explorative group based on the Karapanos et al. (2016)'s classification. Therefore, the participants of this study were predominantly those who started tracking physical activity aiming for healthier lifestyles with their own volition.

Types of data initially and currently monitored with an activity tracker

In the survey, respondents were allowed to select up to three of a total of ten items regarding the type of data they initially wanted to track and currently track, respectively. Over half of participants listed step counts (initial: 65, 57.5%, current: 68, 60.2%), calories burned (initial: 62, 54.9%, current: 62, 54.9%), which was followed by heart rate (initial: 48, 42.5%, current: 41, 36.3%). Activity minutes, distance travelled, and sleep data were also monitored with interest by quite a few people. On the other hand, people showed little interest in flights climbed (initial: 6, 5.3%, current: 5, 4.4%), hours of stand (initial: 2, 1.8%, current: 0). Considering the fact that Apple Watch and Fitbit awards digital badges based on the amount of step counts, calories burned, flights climbed, and hours of stand, this result indicates that current virtual rewards are somewhat misaligned with people's interest. In addition to analysing the frequency of data types, I also examined the change in the data types of interest over time by comparing the responses of each respondent to the two questions. As a result, the interest of the majority of participants (73, 64.6%) shifted from one type of data to another while no change was observed for 40 participants (35.4%).

Frequency of updating the daily goal of activity tracking system

All three devices and apps I investigated in this study encourage their users to set daily activity goals in the initial set up while Apple Watch and Google Fit offer further support by suggesting updating the goals once a week. Since a close connection between the experience

with virtual rewards and goal-setting was suggested from the autoethnographic research, I explored the goal-setting practices of participants. Of the study population, one-third of them (40, 35.4%) rarely update their activity goals. 23 participants (20.4%) reported updating their goals at least once a week, 22 (19.5%) reported once every two weeks, 22 (19.5%) reported once a month, and 6 (5.3%) reported once every two or three months. While approximately 40% of participants pay little attention to their activity goals, a large portion of participants regularly engage in goal-setting practices.

5.2.3. Experiences with Virtual Rewards

Overall attitude, familiarity, and satisfaction

Participants' overall attitudes toward the current virtual reward systems were evaluated by carrying out the content analysis of text responses to the question about an overall impression. The majority of participants (72, 63.7%) had a positive attitude toward badges or points, stating that they benefited from virtual rewards practically or emotionally. On the other hand, 28 (24.8%) deemed virtual rewards pointless while 13 (11.5%) were indifferent to them. When it comes to the familiarity and satisfaction of the current designs of virtual rewards, over half of participants were satisfied with their virtual reward systems (54.9% satisfied or very satisfied, Table 5) while two-thirds of participants showed great familiarity with them (67.3% extremely or very familiar, Table 4). In detail, the overall satisfaction and familiarity between the virtual reward systems were compared using Kruskal-Wallis test since the participant responses did not cover the entire spectrum of the scale in some cases though they were measured on a 5-point Likert scale, which consequently required to perform a non-parametric analysis. A Kruskal-Wallis test showed that there was no statistically significant difference in both overall satisfaction ($\chi^2(2, N=113) = 0.06, p = 0.972$) and familiarity ($\chi^2(2, N=113) = 0.56, p = 0.756$) between the virtual rewards systems. In addition, I examined the degree of satisfaction and familiarity between both age group and usage group. A Mann-Whitney test indicated younger people's satisfaction ($Mdn = 4$) and familiarity ($Mdn = 4$) were not significantly greater than those of older adults' satisfaction ($Mdn = 3$) and familiarity ($Mdn = 4$): satisfaction, $U = 1091, p = 0.073$; familiarity, $U = 1338, p = 0.861$ (Table 6). In terms of usage group, A Mann-Whitney test found no significant differences in the level of satisfaction and familiarity between long-term users (satisfaction: $Mdn = 3$, familiarity: $Mdn = 4$) and short-term users (satisfaction: $Mdn = 4$, familiarity: $Mdn = 4$): satisfaction, $U = 1299.5, p = 0.074$; familiarity, $U = 1499.5, p = 0.566$ (Table 7).

Table 4

Frequencies of Familiarity of the current badge/points system

	Apple Watch	Fitbit	Google	Total	Percent
Not at all familiar	2	5		7	6.2%
Not so familiar	2	5	1	8	7.1%
Somewhat familiar	7	12	3	22	19.5%
Very familiar	13	22	5	40	35.4%
Extremely familiar	8	23	5	36	31.9%
	32	67	14	113	100.0%

Table 5

Frequencies of Satisfaction of the current badge/points system

	Apple Watch	Fitbit	Google	Total	Percent
Very dissatisfied	2			2	1.8%
Dissatisfied	3	3		6	5.3%
Neither satisfied nor dissatisfied	9	27	7	43	38.1%
Satisfied	8	20	4	32	28.3%
Very satisfied	10	17	3	30	26.5%
	32	67	14	113	100.0%

Table 6

Familiarity and satisfaction between age groups

	Younger (18-34)		Older (35 or more)		U	Z	p
	Mean Rank	n	Mean Rank	n			
Familiarity	56.65	78	57.77	35	1338	-0.18	0.861
Satisfaction	60.51	78	49.17	35	1091	-1.79	0.073

Table 7

Familiarity and satisfaction between usage groups

	Short-term (6 months or less)		Long-term (7 months or more)		U	Z	p
	Mean Rank	n	Mean Rank	n			
Familiarity	55.35	58	58.74	55	1499.5	-0.57	0.566
Satisfaction	62.09	58	51.63	55	1299.5	-1.79	0.074

Lived experiences with virtual rewards

Participants reported one or more notable experiences that they were positively or negatively influenced by virtual rewards in their physical activity practices. The experiences of 106 participants were investigated through the content analysis with 7 out of the study population excluded due to the lack of detail. Since some answers included multiple experiences which were handled as an individual account, 194 accounts in total were generated and coded. Then, 23 codes were organized into 16 categories, among which 5 are about psychological effect, 7 are about practical use, and 4 are about negative attitude and influence. Code frequencies of each category were calculated and presented in Table 8.

Psychological benefit

Though the sample size of Google Fit, which represents a point-based reward system in this study, is much smaller than that of other two badge systems, psychological benefits were mostly observed in the user group with digital badge system. 22 participants explicitly stated that achieving virtual rewards gave them a sense of pleasure, and a digital badge system added additional playfulness to activity tracking practices as described by one participant:

“Reaching our goals gives us an energetic and fresh feel. This badging system gives me pleasure of using the tracker.” (P7, Apple Watch)

Also, many participants mentioned they felt a sense of achievement when they earned badges. The results indicated that this mainly comes from 1) accomplishing activity goals, 2) succeeding challenging tasks, such as sticking to a routine, breaking the record, or doing an enormous amount of workout, and 3) appreciating the collection of badges:

“I feel accomplished when I hit one of the big goals. The single dailies don’t matter too much but when I hit a monthly, 300-500% move goal, all three in a day, or major milestone, I legitimately feel accomplished and it makes me want to work harder for more.” (P26, Apple Watch)

“When I first got the Marathon badge for completing 26 miles. I was so excited that I informed my whole family about it as I was a very lazy person who wouldn’t get off from his computer for anything other than food.” (P56, Fitbit)

For some people, earning badges and points made them feel better about themselves:

“Once I realized I was able to walk so many steps in one day it was really surprising, and I figured that I could do more steps if I spend more time pursuing it. My first indicator of such a feat was Boat Shoes.” (P92, Fitbit)

Positive effects on physical activity practices

In addition to psychological benefits, some utilitarian benefits were also confirmed regardless of the type of virtual rewards. While Both Apple and Fitbit’s badge systems seemed to make

Table 8

Code frequencies of 16 categories identified

Themes	Categories	Codes (Experiences)	Frequency		
			Badge	Point	Total
Hedonic	Pleasure	I felt a sense of pleasure	21	1	22
	Achievement	I felt a sense of achievement	21		21
	Pride	I felt a sense of pride (felt good about myself)	8	1	9
	Surprise	It was an unexpected reward	5		5
	Aesthetic	I was influenced by aesthetic	2		2
Utilitarian	Increased Engagement	I was motivated to exercise more	14	3	17
	Challenging	I completed challenging goals (major goal, streak, best record)	15		15
	Increased Enjoyment	I was motivated to earn more badges (badging experience)	15		15
	Progress	I confirmed progress (milestones)	11		11
	Increased Engagement	I was encouraged to stay on track	10	1	11
	Alternative Goal	I targeted them as an alternative goal	9		9
	Progress	I felt appreciated (I felt like I'm on the right track)	5	3	8
	Informative	It deepened my understanding of fitness routine	2	3	5
	Memorial	It was the first badge I earned	5		5
	Increased Enjoyment	It gamified the experience (gamification)	2	1	3
	Memorial	It came after the actual health improvement	3		3
Negative	Indifferent	I am Indifferent	9	3	12
	Pointless	It is pointless	8		8
	Insufficient Credits	I was discouraged by the lower achievements than expected	4	1	5
	Novelty effect	The novelty effect wore off	3		3
	Interval	I was discouraged by the interval getting longer	2		2
	Decreased Interest	I was more motivated before	2		2
	Unrealistic	I was discouraged by seeing the unreachable badges	1		1
Total			177	17	194

physical activity more enjoyable, the Google's point system offered greater informational value to users in relation to their health and daily practices. Many of those who had used digital badge systems were attracted to badging experience itself. They liked unlocking and collecting badges, which made them motivated to exercise more and continue to engage in physical activity:

"I remember that I was falling behind on a monthly challenge badge where I had to walk a certain distance to acquire the badge. It was the last day of the month, so I forced myself to get up and I began to walk the last few miles to achieve the badge. I felt like this was extremely positive." (P25, Apple Watch)

"The first time I achieved the badge showing how many miles I walked, I felt so proud and more motivated to continue getting at least that number, as well as trying to get more." (P91, Fitbit)

As reported in the previous studies (Fritz et al., 2014; Vooris et al., 2019), some participants did extra workout just for achievable badges when they realized their activity levels were close to the thresholds for those badges:

"I have been close to a goal or badge and since I know I can achieve that I kept taking steps to get it." (P90, Fitbit)

For some people, their desire for digital badges led to trying different types of exercises, resulting in a change of their attitude towards physical exercise:

"The first badge that I received was the boat shoes, I didn't realize how much receiving a badge would make me want to get another badge. Staying active has become a positive challenge instead of chore/hassle that I want to simply get out of the way. I look forward to yearning new badges." (P39, Fitbit)

"I was positively influenced by badges when I earned the cycling badge. I despise cycling but the motivation to earn that badge made me do some cycling and when I earned that badge it made it seem much more worthwhile." (P3, Apple Watch)

Speaking of more practical contribution of virtual rewards to people's everyday practices, participants described that digital badges served as an alternative activity goal which they could aim at:

"I really like it because it makes me feel like I have something to work towards." (P40, Fitbit)

"It gives me a goal to obtain. It makes me feel like being healthy is equivalent to winning a game." (P39, Fitbit)

Also, participants confirmed that they were making good progress on their fitness routines through the badges or points they earned:

"Regularly I am positively influenced by these points when I realize that I earned a lot more than I expected." (P106, Google Fit)

“It’s cool, it creates a nice simple way of feeling like I’ve crossed into a new stage or met a persistent goal and usually is just a simple way of tracking my goals and feeling like I’ve met them.” (P53, Fitbit)

Furthermore, badges and points could help people gain a deeper understanding of their activity level and exercise routines:

“I was positively influenced by digital badges when I first learned about it. It actually helped me form a schedule and see how my routine could collect awards.” (P4, Apple Watch)

“I once checked my move minutes and was shocked about how sedentary I was. It made me start to walk more.” (P110, Google Fit)

Negative effects

As suggested by previous study (Antin & Churchill, 2011) and revealed through their overall impression, not all participants appreciated virtual rewards. They seemed like a background noise for those who were indifferent to badges or points:

“Honestly I can’t think of either. I just never cared, but I would say sometimes it was invasive when I just wanted to check there would be a reminder that I have completed a new badge. It kind of annoyed me. I don’t care.” (P73, Fitbit)

Moreover, some participants pointed out the misalignment between their health-related concerns and the current designs of virtual rewards:

“As I said, the badges are irrelevant to me. My health is what matters to me, and my fitness progress is what I have measured. This isn’t a video game to me.” (P98, Fitbit)

Meanwhile, the novelty effect of digital badges was also revealed by some participants. This could be caused by many reasons, but the lack of practical value and the gap between current virtual rewards and traditional reward experiences were mainly pointed out:

“They never really meant much to me. I was excited when i got them the first time, but they don’t really show anything.” (P84, Fitbit)

“I find the experience neutral to negative because the different shoes don’t really translate into anything I traditionally recognized as a reward system.” (P54, Fitbit)

Also, some participants reported that they were discouraged or lose their interest due to insufficient credits for their exercise, a limited number of badges, long time intervals between rewards, and unrealistic thresholds:

“During the first couple of weeks I got a lot of badges, I guess it’s just supposed to be like that - encourage people, so that they don’t lose interest. But the more you wear your device - the less badges you get. It can be months without a badge. This feels bad. Not really encouraging.” (P48, Fitbit)

“Some actions you would think you did longer or worked harder, but some seem to be quite similar. Such as walking or brisk walking.” (P105, Google Fit)

Valuable and unimportant badges or points

Based on the wearable device and its dedicated app they are currently using, participants reported both the most and the least appreciated or valuable badge(s)/point(s) with reasons. Each response was validated based on the knowledge gained from desk research and autoethnography, as a result of which some responses with the lack of detail were excluded for the analysis in each case.

Digital badge system

When it comes to valuable badges, both daily Move Goal badges and Perfect Week badge were named by 6 participants (21.4%) for Apple Watch (N=28) while daily steps badges were chosen by 42 participants (67.7%) and lifetime distance badges by 12 (19.4%) for Fitbit (N=62). To be specific, the most frequently named Fitbit's daily step badges were Sneakers (9), followed by Urban Boots (6), Hiking Boots (5), High Tops (5). Users can earn these badges by completing 10,000 steps, 15,000 steps, 35,000 steps, and 20,000 steps, respectively. The reasons provided by participants gave an explanation of their choices. Participants chose these badges because these badges are challenging and show their enormous effort:

“It showed how committed I was.” (P21, Apple Watch)

“I climbed a really high mountains and I like having a little badge of honor about it.” (P62, Fitbit)

Also, participants deemed lifetime distance badges informative in that they show how many miles they have walked since they started using the device:

“To know that I have walked the diameter of the earth will be a great accomplishment. It lets me know that if I could truly walk around the world, I could make it.” (P39, Fitbit)

Regarding unimportant badges, First Workout badges (8, 33.3%) was top on the list, followed by Activity Challenge badges (6, 25.0%) for Apple Watch (N=24). The most frequently mentioned badge by Fitbit users (N=58) was one of daily steps badges (34, 58.6%), among which two-thirds (23, 67.4%) stated Boat Shoes which is awarded for completing 5,000 steps. The second highest one was one of daily floors badges (12, 20.7%). Of those 12 participants, 7 (58.3%) mentioned Happy Hill, which is given for the completion of 10 floors. Both Boat Shoes and Happy Hill are the easiest badges on each badge list. When a badge is too easy to get, people find it pointless:

“I felt like it was given out for free.” (P15, Apple Watch)

“I can get 5,000 steps a day with my eyes closed.” (P33, Fitbit)

Also, people regarded a badge meaningless when it is not aligned with their interests:

"I'm not at all interested in yoga." (P3, Apple Watch)

"I am not a fan of climbing and so it doesn't matter to me much." (P63, Fitbit)

Points system

Compared to the badge systems of Apple Watch and Fitbit, Google's points system has only two options, Move Minutes and Heart Points. Thus, participants (N=14) were asked to select the point(s) that they deem valuable from one of four options: Both, Heart Point, Move Minutes, and Neither. 6 out of 14 participants (42.9%) chose both, followed by Heart Points (3, 21.4%). 4 (28.6%) reported that neither of them was valuable to them. Those who chose at least one of the points stated that points are a good indicator of their activity level and useful for maintaining their health:

"I believe both are vital to one's general well-being." (P101, Google Fit)

"Move minutes shows the amount of time I was moving around. It helps to monitor my activity level." (P107, Google Fit)

Those who selected none of them generally had no interest in virtual reward systems:

"I just want to track my workouts for me, to make sure I am hitting MY goals. I don't care about points or what Google's AI thinks about my workouts." (P113, Google Fit)

Sharing Experience

Participants answered yes or no to the question of whether they had ever shared a virtual reward. As a result, more than three quarters of participants (87, 77.0%) had never shared their rewards with others. Most of them do not feel a need to show their achievements to others:

"I just keep my accomplishments to myself. No sense in telling everyone." (P3, Apple Watch)

The reason behind this seemed that participants regarded activity- or health- tracking as a personal thing:

"I don't want to show others what I have done because I want my fitness level to be improved." (P7, Apple Watch)

"I have always felt my health was a personal thing." (P110, Google Fit)

Participants did not think people around them would care even if they wanted to share:

"I just don't have anyone around me that is working out so I don't share anything about my badges because I feel most people wouldn't care about it." (P75, Fitbit)

On the other hand, those who answered yes were primarily sharing virtual rewards to share

their progress with friends, family, and other users or to help others get motivated:

“My wife has the same Fitbit, so from time to time we discuss the badges - it’s some kind of competition for us.” (P48, Fitbit)

“I share my badges with others who use fit bit. It’s a nice and easy way to share motivation with another.” (P39, Fitbit)

Perception and behavioral and psychological effect

Overall interpretation based on descriptive statistics

Table 9 shows the results of Likert scale items. Since one question measuring a detrimental effect of virtual rewards is a negative item, values of that item were reversed before interpreting the result (abbreviated as E_Negative_r in Table 3), by which the higher the score, the more positive. Given that the midpoint of the Likert scale item was 4 which indicates neither agree nor disagree, participants’ responses were somewhat positive in all three categories: quality of system ($N=99$, $M=5.20$, $SD=1.06$), perception and attitude ($N=113$, $M=5.00$, $SD=1.57$), and perceived effect ($N=113$, $M=4.84$, $SD=1.38$). When looking closely at the individual items, the mean of all the items were above the midpoint while some items scored relatively lower or higher than the mean of the corresponding categories. When it comes to the perceived value of virtual rewards ($N=113$, $Mdn=5$), more than half of the participants (57.5% somewhat agree, agree or strongly agree) found badges/points valuable while around half of people still did not, which is relatively low compared to the proportion of those who answered positively in other categories. In the same way, it was assumed that perceived social effect of virtual rewards ($N=113$, $Mdn=4$) was also weak: 47.8% expressed agreement while 52.2% did not. On the other hand, the results suggested that people had no difficulty in understanding how virtual reward system works ($N=113$, $Mdn=6$) as well as in achieving virtual rewards ($N=113$, $Mdn=6$). The proportions of those who agreed with the statement were 84.1% and 82.3%, respectively.

Validity and reliability of Likert scale questionnaire

Before carrying out any inferential analysis, I conducted a factor analysis with the maximum likelihood method with varimax rotation to test the construct validity of a Likert scale questionnaire. Two items about the frequency and diversity of rewards were excluded since they were only answered by those who use activity tracking systems with digital badges. As a result of the factor analysis with 22 Likert scale items, two factors were detected to explain 71.4% of the total variance (Table 10). While 3 out of 22 items were removed since both of two items loaded almost equally on two factors and another item did not have enough factor loadings on either of factors, 19 items fell into two groups. One factor consisted of 4

Table 9

*Mean, standard deviation, and percentage of frequencies of Likert scale items (N=113), *N=99*

Items	Mean	SD	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
S_Aesthetic	5.08	1.38	3.5%	1.8%	4.4%	17.7%	32.7%	25.7%	14.2%
S_Access	5.46	1.41	0.9%	2.7%	6.2%	15.0%	18.6%	28.3%	28.3%
S_Ease	5.67	1.31	0.9%	0.9%	6.2%	8.0%	23.9%	26.5%	33.6%
S_Trust	5.27	1.56	2.7%	3.5%	6.2%	18.6%	15.9%	27.4%	25.7%
S_Support	5.22	1.64	5.3%	2.7%	5.3%	15.0%	19.5%	27.4%	24.8%
S_Frequency *	4.46	1.89	8.1%	10.1%	14.1%	17.2%	13.1%	20.2%	17.2%
S_Diversity *	5.27	1.43	2.0%	2.0%	8.1%	14.1%	22.2%	31.3%	20.2%
P_Emotion	5.2	1.82	7.1%	5.3%	4.4%	10.6%	17.7%	25.7%	29.2%
P_Value	4.58	2.00	11.5%	9.7%	7.1%	14.2%	15.9%	22.1%	19.5%
P_Credit	4.95	1.77	7.1%	3.5%	9.7%	15.9%	15.0%	28.3%	20.4%
P_Enjoy	4.97	1.87	9.7%	3.5%	8.0%	9.7%	19.5%	26.5%	23.0%
P_Difficult	5.57	1.34	2.7%	0.9%	1.8%	12.4%	23.9%	31.0%	27.4%
P_Curiosity	4.66	1.88	8.8%	9.7%	8.0%	11.5%	19.5%	26.5%	15.9%
P_Willingness	5.09	1.90	8.0%	6.2%	6.2%	9.7%	17.7%	23.0%	29.2%
E_Goal	4.96	1.81	6.2%	7.1%	8.8%	9.7%	21.2%	23.9%	23.0%
E_Progress	5.4	1.70	5.3%	3.5%	4.4%	11.5%	15.0%	28.3%	31.9%
E_Motive	5.16	1.87	8.0%	5.3%	5.3%	10.6%	14.2%	28.3%	28.3%
E_Pride	5.02	1.83	7.1%	6.2%	8.0%	9.7%	17.7%	28.3%	23.0%
E_Feedback	4.9	1.71	5.3%	3.5%	12.4%	17.7%	16.8%	23.9%	20.4%
E_Enjoy	4.83	1.90	8.8%	7.1%	8.0%	9.7%	24.8%	17.7%	23.9%
E_Effort	4.94	1.98	10.6%	5.3%	8.8%	7.1%	16.8%	25.7%	25.7%
E_Social	4.09	2.08	16.8%	12.4%	9.7%	13.3%	18.6%	12.4%	16.8%
E_Competetion	4.72	1.93	9.7%	8.0%	10.6%	6.2%	22.1%	24.8%	18.6%
E_Negative_r	4.42	2.07	7.1%	18.6%	15.9%	5.3%	13.3%	16.8%	23.0%

Table 10

Final constructs after factor analysis

Final Constructs	Initial Constructs	Items	Cronbach's alpha	Composite item variable	
				Mean	SD
System Quality	Quality of System	S_Ease, S_Trust, S_Access, P_Difficult	0.76	5.49	1.08
Perception and Effect	Perception and Attitude / Behavioral and Psychological Effect	P_Emotion, P_Enjoy, P_Value, P_Credit, P_Willingness, P_Curiosity, E_Competetion, E_Effort, E_Pride, E_Enjoy, E_Motive, E_Goal, E_Progress, E_Feedback, E_Social	0.98	4.90	1.64

items related to system usability and credibility, which I will call “System Quality” construct. The other factor with 15 items, which accounts for 62.57% of the total variance, will be referred to as “Perception and Effect” construct since it was a combination of items originally designed to measure attitudes and influences. To evaluate the internal consistency of items in each construct, Cronbach’s alpha was calculated. Since the coefficients for both constructs exceeded 0.7, which is generally known as an acceptable level, two constructs were found to be sufficiently reliable. Based on the results above, the mean score of each construct was calculated using sub-items and passed to the further analysis.

Comparison of System Quality and Perception and Effects between user groups

To examine whether there is a significant difference between virtual reward systems in terms of System Quality and Perception and Effect, A One-way Analysis of Variance (ANOVA) was performed using the average score of two constructs. Table 11 indicates the means and standard deviations for each of the three groups. Since the sample sizes for each group are quite different and one of them is particularly small ($n=14$), the *Welch’s F* test was used. The result showed that there was no statistically significant difference in the quality of system and users’ attitudes and perceived effects between groups (Table 12): System Quality, *Welch’s F* (2, 34.98) = 0.33, $p = 0.72$; Perception and Effect, *Welch’s F* (2, 33.24) = 0.58, $p = 0.56$. As with satisfaction and familiarity, I evaluated whether there was a difference between the two scores by age group and usage group. *Welch’s t*-test indicated the scores of younger group were not significantly higher than those of older group in terms of both System Quality and Perception and Effect (Table 13): System Quality, t (57.80) = 1.57, $p = 0.123$; Perception and Effect, t (52.34) = 1.63, $p = 0.11$. Also, an independent-samples *t*-test found that there was no significant difference between long-term and short-term user group (Table 14): System Quality, t (105.52) = 0.22, $p = 0.827$; Perception and Effect, t (110.14) = 0.80, $p = 0.427$.

Table 11

Mean and standard deviation of the mean scores of System Quality and Perception and Effect by virtual reward system

Virtual Reward System	<i>n</i>	System Quality		Perception and Effect	
		<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
Apple Watch	32	5.39	1.10	5.13	1.30
Fitbit	67	5.50	1.09	4.80	1.72
Google Fit	14	5.66	1.01	4.82	2.02

Table 12

One-way Analysis of Variance of system Quality and Perception and Effect by virtual reward system

Constructs	Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>
System Quality	Between Groups	2	0.74	0.37	0.33	0.72
	Within Groups	34.98	129.25	1.18		
	Total	112	129.99			
Perception and Effects	Between Groups	2	2.48	1.24	0.58	0.56
	Within Groups	33.24	300.32	2.73		
	Total	112	302.80			

Table 13

Welch's t-test for System Quality and Perceived Effects (Younger vs. Older)

	Younger (18-34, <i>n</i> =78)		Older (35 or more, <i>n</i> =35)		<i>df</i>	<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
System Quality	5.60	1.02	5.24	1.18	57.80	1.57	0.123
Perceived Effects	5.08	1.47	4.49	1.94	52.34	1.63	0.11

Table 14

Independent Samples t-test for System Quality and Perceived Effects (Long-term vs. Short-term)

	Short-term (6 months or less, <i>n</i> = 58)		Long-term (7 months or more, <i>n</i> = 55)		<i>df</i>	<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
System Quality	5.51	0.99	5.47	1.18	105.52	0.22	0.827
Perceived Effects	5.02	1.62	4.77	1.68	110.14	0.80	0.427

Association between satisfaction and individual items

To estimate the association between satisfaction/familiarity level and each of the 24 Likert scale items, a Kendall's tau-b correlation was carried out. In terms of satisfaction (Table 15), the correlations were mostly positive and statistically significant (Kendall's tau = 0.28-0.54, $p < 0.01$), except one to measure the degree of obsession with virtual rewards (Kendall's tau = -0.20, $p = 0.09$). In particular, a strong, positive correlation was estimated between satisfaction level and motivational effects: E_Motive, Kendall's tau = 0.50; E_Effort, Kendall's tau = 0.54. In other words, if a participant had been motivated for increased engagement with

physical exercise through virtual rewards, that person had been more likely to be satisfied with the virtual reward system. Regarding familiarity (Table 16), the correlations were weak and statistically significant for only some items, among which there was a moderate, positive correlation between ease of use and familiarity level (Kendall's tau = 0.34, $p < 0.01$). That is, if a respondent had easily figured out how to engage in the virtual reward system, that person might have felt more familiar with the system.

Table 15

Correlation between satisfaction and individual Likert scale items

System Quality	S_ Aesthetic	S_ Access	S_ Ease	S_ Trust	S_ Support	S_ Freq	S_ Diverse			
Correlation (τb)	.359**	.278**	.364**	.465**	.412**	.360**	.361**			
Perception and Attitude	P_ Emotion	P_ Value	P_ Acknowl	P_ Enjoy	P_ Difficult	P_ Curious	P_ Will			
Correlation (τb)	.495**	.438**	.484**	.496**	.347**	.465**	.491**			
Behavioral and Psychological Effects	E_ Goal	E_ Progress	E_ Motive	E_ Self	E_ Feedback	E_ Enjoy	E_ Effort	E_ Social	E_ Compet	E_ Negative_r
Correlation (τb)	.467**	.454**	.502**	.420**	.283**	.460**	.535**	.318**	.412**	-.201**

* $p < .05$, ** $p < .01$

Table 16

Correlation between familiarity and individual Likert scale items

System Quality	S_ Aesthetic	S_ Access	S_ Ease	S_ Trust	S_ Support	S_ Freq	S_ Diverse			
Correlation (τb)	.225**	.224**	.341**	.313**	.133	.157	.04			
Perception and Attitude	P_ Emotion	P_ Value	P_ Acknowl	P_ Enjoy	P_ Difficult	P_ Curious	P_ Will			
Correlation (τb)	.220**	.228**	.208**	.236**	.230**	.209**	.198*			
Behavioral and Psychological Effects	E_ Goal	E_ Progress	E_ Motive	E_ Self	E_ Feedback	E_ Enjoy	E_ Effort	E_ Social	E_ Compet	E_ Negative_r
Correlation (τb)	.140	.146	.159*	.138	.119	.125	.155*	.104	.242**	.007

* $p < .05$, ** $p < .01$

Suggestions for the improvement of virtual reward system

At the end of the survey, I asked participants about how the current designs of virtual reward system can be improved. Through a content analysis of the responses from 70 participants, a total of 84 needs and suggestions were collected. Among which, the need for practical value of reward (24) was greatest, including needs for physical rewards (14) and monetary benefit (10):

“I’m pleased with it as it is. The only improvement I could see is if certain badges came with real-life rewards (like maybe you get a new Fitbit for free if you collect a certain very high number of badges).” (P77, Fitbit)

“Maybe entering high achievers into some sort of sweepstakes to win actual prizes.” (P74, Fitbit)

Some wanted to add more meaning to the rewards (9) or customize a badge to fit their activity goals (6):

“Make them mean something, or give you access to premium features or something.” (P72, Fitbit)

“Specific customization as to what I want those goals to be, such as naming them myself.” (P29, Apple Watch)

Along with the need for practical use, another great need (16) was a more diversified badging experience by having a greater variety of badges or by having different badges or upgrading badges based on their fitness improvement:

“I think if they had many options for the same badge like different graphics or something like that, it’d be fun.” (P82, Fitbit)

“The badges should be upgraded with next motivational goals.” (P51, Fitbit)

Also, a need for aesthetic improvement was also expressed by some (8):

“Maybe make them animated? Like a gif or something.” (P2, Apple Watch)

5.3. Summary of Results

Through an online survey with 113 activity tracker users, the situated experiences of users were vividly drawn, and diverse aspects of virtual reward experience were qualitatively and quantitatively explained. First, it was identified that the majority of the survey participants took a positive attitude toward virtual rewards. By analyzing the narratives of their experiences, both hedonic and utilitarian benefits of virtual rewards became clearer. In addition, several qualitative questions revealed the personal nature of activity tracking and clarified why users add more value to certain types of virtual rewards. Meanwhile, a quantitative investigation was also attempted through Likert scale items, as a result of which the positive correlation between the level of satisfaction and motivational effects was confirmed. Interestingly, people tended to be more obsessive about virtual rewards as their satisfaction increased. In addition, with inferential statistics, no statistical differences in satisfaction and perceived effects of virtual rewards were found between the virtual reward systems, age groups (younger vs. older), and usage groups (long-term vs. short-term) although the average scores of younger adults and short-term users were higher than their counterparts. Lastly, by collecting people's suggestions, a need was identified for improved practicality and meaning of virtual rewards and a more diversified, long-lasting virtual reward system.

Discussion

In this chapter, I present an interpretation of the results to answer my research questions. The most important findings are discussed, as well as the implications for designing a virtual reward system for promoting physical activity.

6.1. Key Findings

This work provided a more nuanced understanding of how people perceive and experience current virtual reward systems in daily practices. The study revealed several things that were not studied in the prior research while lending support to previous findings in the literature.

Effects of virtual rewards in everyday practices

The reviewed literature pointed to the mixed views on virtual reward systems in activity tracking systems. Some studies revealed motivational impacts of virtual rewards on people's physical activity practices, whereas other studies reported that the effectiveness of virtual rewards were limited to a few users while many people were indifferent to their virtual reward systems. While these findings were drawn from a limited number of participants in qualitative research, the results of this study derived from 113 participants who were recruited from an online crowdsourcing platform without any personal connection with the researcher. In this sense, this study could complement the prior studies by providing further evidence. First of all, the results of the survey with 113 actual users indicated that two-thirds of participants had a positive attitude or benefited from the current virtual reward. Moreover, some psychological and behavioral benefits of virtual rewards in daily practices were also revealed. Participants reported that collecting virtual rewards motivated them to exercise more and helped them stay active while giving a sense of pleasure and achievement. Also, badges could serve as an alternative activity goal, sometimes putting them on an extra workout to get them. These are consistent with the findings in the reviewed literature. Furthermore, the study provided new evidence on the informational roles of digital rewards. By earning rewards from the systems, some participants had more confidence in what they were doing or deepened their understanding of the current activity level and fitness routines.

What makes some virtual rewards valuable?

Antin and Churchill (2011) indicated through their research on Foursquare that most people find hedonic or behavioral value from only some types of badges. This was also confirmed in both autoethnographic research and online survey, with similarities between the results. I found greater value from virtual rewards when they are 1) recognizing my effort and time appropriately, 2) informing the impact of my exercise on health, 3) challenging to achieve, 4) new, and 5) aligned with my interest and context. Similar results were gained by analysing the participants' responses for both the most and the least appreciated or valuable badge(s)/point(s) with reasons. The results suggested that the value of virtual rewards mainly derives from when they are 1) providing proof of enormous efforts, 2) challenging to users, 3) informative, 4) giving assurance of users' activity level, and 5) helping to engage in physical

activity. On the other hand, the virtual rewards may fail to yield intended outcomes when they are 1) not interesting and meaningful to users, 2) too easy to achieve, 3) offered repeatedly, 4) not suitable for users' ability or context. By comparing and combining the results from both autoethnography and online survey, it became evident what makes virtual rewards valuable or less important in the context of physical activity tracking.

Personal nature of activity tracking and virtual rewards

Social functionality is often employed (Zuckerman & Gal-Oz, 2014) in systems designed for encouraging physical activity. Among three activity tracking systems studied in this work, Fitbit and Apple Watch offer social functions to users. However, the study revealed that there exists a gap between the expectation of the system designer and actual use in terms of social functionality. The majority of the study population had never shared virtual rewards with others. The reasons provided by participants can give an explanation of this gap. Since users considered activity- or health- tracking as a personal matter, they did not want to share their accomplishments, nor did they feel the need to share them with others. Even if they wanted to, they didn't think people around them would care. Given that a badge's value often comes from a community that put value upon that badge (Cugelman, 2013), this personal nature of activity tracking poses a challenge to the designers who develop a virtual reward system, especially a digital badge system, within this context.

Virtual rewards as part of engagement loop

From the diary-based autoethnography, how the current virtual reward systems interact with both goal-setting and notification features was illustrated in detail. When the rewards were awarded based on the attainment of daily activity goal like Apple Watch in this study, setting a doable but challenging goal was important to have a better rewarding experience through virtual rewards. Also, as with Google Fit, the experience of earning virtual rewards could be reinforced by accomplishing some health-related milestones. In terms of notification, informing the progress of rewards could keep people motivated to achieve the rewards and engaged in physical activity. Moreover, it seemed more effective when rewards were delivered on the spot in real time. Therefore, supporting goal-setting practices and communicating properly with users can make virtual reward more an effective tool. This close connection between the system features can be explained by the concept of social engagement loop. A social engagement loop is designed to increase and sustain user engagement, in which each of the four elements—motivating emotion, social call to action, player re-engagement, and visible progress/reward—leads to the next one and return to a motivating emotion again (Zichermann & Cunningham, 2011). Applying this concept to the study context, motivating emotion can be achieved through setting an appropriate level of activity goal

while notifications play a pivotal role in social call to actions and player re-engagement in this loop. Thus, careful consideration of how a virtual reward system interlocks with other features such as goal-setting and notification is required since virtual rewards perform not as the sole motivator, but as a part of the engagement loop.

More complete image through the use of mixed methods

To answer the research questions, two entirely different approaches were taken: a diary-based autoethnography and an online survey. Although one is inherently subjective and qualitative while the other is more objective and quantitative, the results were found to be complementary to each other. A total of four months of autoethnography listed in detail many aspects of the virtual reward experience that were difficult to find in surveys or interviews. For example, without firsthand experience and the increased awareness of the present moment, it would have been difficult to reveal how a virtual reward system interacts with other features and how those interactions could affect users' behaviors and attitudes. On the other hand, an online survey provided a lot of stories about users' situated experiences in a short time. Also, the quantitative analysis more clearly showed users' behaviors and attitudes with numbers while compensating the results of autoethnography. This enabled me to elicit the true nature of sharing experience around virtual rewards and the factors that make some types of rewards more valuable and meaningful. All in all, using both qualitative and quantitative methods facilitate a more robust and empathetic understanding of users and their experiences.

6.2. Design Considerations

While this study detailed the real-life experience with virtual reward systems in activity trackers, the findings from the literature review and user studies allowed me to discuss several considerations for the design of more effective virtual reward system within this context.

6.2.1. User-Centered Reward System

Aesthetics

Among the study population, there was a need for aesthetic improvement of current designs of badges or the way of visualizing accrued points. On the other hand, a few participants were attracted to the color or animation of badges. Since the preferences for visual design vary from person to person, it is hard to generalize which design is superior or inferior than others. Some prefer more realistic images whereas others prefer illustrations. Thus, I would like to focus on the animation and interactivity of badges here. Compared to physical rewards, digital rewards have some limitations since they are intangible. However, virtual rewards can have an advantage in that they can be more dynamic and interactive with users. One way of taking advantage of digital object is flipping motion and allowing users to fiddle with badges that are applied to Apple Watch's badges. It can be approached in many ways, and by doing so, digital badges could have their own unique value beyond just a visual representation of achievement.

Accessibility to reward system

The autoethnographic diary study identified that the current activity tracking systems provide different levels of accessibility to their virtual reward systems and that easy access to virtual rewards leads to higher awareness of achievements. Also, the survey participants described that they felt a sense of achievement by appreciating their collections of digital badges and that they confirmed they were making good progress through the badges they had earned since they started, so-called, their fitness journey. Since the study results indicated that easy access to reward system could lead to the higher level of awareness as well as a good sense of achievement and self-affirmation, higher accessibility to both individual reward and the reward system needs to be provided, such as by allowing alternative routes to access them or by bringing them to the top of the hierarchy.

Transparency of system

Nicholson (2012) pointed out that it is critical to make systems more transparent by providing appropriate justification to support users' decision-making process and giving sufficient information of what is going on in order for user-centered meaningful gamification. In my experience with three virtual reward systems, there were moments I wondered why in all cases. Each time, I examined the reasons myself or through online searches. Interestingly, when searching online, I was always able to find someone who already had the same question, which was answered by another user. When I experienced the inconsistency of system without justification, I felt discouraged and lost my trust in the reward systems. For these reasons, it should be reviewed whether the reward system provides users with adequate information about the requirements for a reward as well as appropriate justification for their success or failure to achieve the reward.

6.2.2. Multi-Layered Virtual Reward Experience

Deterding (2013) stated that existing gamified systems merely added game design elements and limited themselves to a small set of feedback interface design patterns. This statement was made a couple of years ago, but it still seems valid to the virtual reward system in current activity trackers considering the evidence found in this study. Some participants described that they lost their interest in a digital badge system over time due to a limited number of badges and long-time intervals between rewards or that they were discouraged by unrealistic requirements for some badges. Furthermore, among the responses to the suggestion for improvement, many participants wanted to have a greater variety of badges. Also, some suggested applying a level system to the current badge systems, by which available badges are also going to be upgraded with their fitness levels or goals. High achievers wanted to have some special benefits while low achievers wanted to have easier goals. In sum, all these comments pointed to a need for more diversified virtual reward experiences. Given that users' activity goals and ways of using activity trackers have evolved over time (Fritz et al., 2014; Gouveia et al., 2015), this kind of approaches to improving the design of virtual reward systems makes sense. Also, it seems worthwhile to consider assigning badges a different role. Some of the participants regarded a badge as an alternative goal for their daily activity while some deemed a challenging badge as a major milestone in their fitness journey. Thus, providing both options in a separate manner with different logics can be proposed. For example, a reward system could offer a daily goal badge in return for achieving a certain number of manually set goals while allowing users to start challenging goals when they are ready. This suggestion could also deepen and diversify the experience with virtual rewards.

6.2.3. Meaningful Reward Experience

Increased informational value with health-related information

Several findings pointed to a need for more practical use of virtual rewards. Some survey participants reported that they were indifferent to virtual rewards due to the lack of perceived practical value. Also, the score of the Likert-scale item measuring perceived value was relatively lower than others. Furthermore, a need for practical and meaningful use was the most frequently found among the suggestions. Though many of them wanted to find monetary value from badges or points, monetary incentives may have a detrimental effect on those who have low self-efficacy (Tseng, Chang, & Yen, 2018). In relation to this, Tseng et al. (2018) suggested monetary incentives may be more beneficial to the goal commitment after users have self-set goals and increased engagement in physical exercise. Therefore, as another way to heighten practical value, providing increased informational value seems most effective and reasonable. This can be backed up by the findings from the reviewed literature, autoethnography, and online survey. Nicholson (2012) claimed that providing relevant information to users' interests helps people find meaningful connections between game design elements, activity, and their goals, which results in meaningful gamified experience. In the context of using activity trackers, the primary purpose of adopting the device in everyday life is managing and improving health as proved in the survey as well. So, the most relevant information to user's interest in this context is health-related information. Moreover, in my experience with Google Fit, I appreciated the informational value of Heart Points. I internally linked the amount of Heart Points to the impact of exercise on my health. Also, some survey participants criticized the misalignment between their interest in health and the current content of rewards. In this sense, Google Fit seems to be a major step in the right direction although its effectiveness still needs to be verified with a large population.

Personalization/Customization

Another way to enhance perceived value of virtual reward system is adapting the system to changing personal interests and needs. Fritz et al. (2014) and Gouveia et al. (2015) indicated that the activity goal of users and their engagement with activity trackers have continuously changed as time goes on. A similar trend was observed in the survey as well. By comparing the types of data initially and currently monitored with an activity tracker, the interest of two-thirds of participants changed from one type of data to another. In addition, this survey question revealed that there is a misalignment to some extent between users' interests and the data types of badges available. Based on my experiences with three different virtual reward systems, this misalignment means a lot. When I was not interested in the data types of badge, I became indifferent to that, and sometimes, earning those badges was like a background

noise. To address this, some participants suggested creating badges tailored to their activity goals. Meanwhile, a recent study on adaptive interventions for increasing physical activity showed the potential of this approach. Korinek et al. (2018) conducted a 14-week field trial with 20 participants to evaluate an adaptive step goal and rewards intervention using a smartphone app, in which users' step goals were personalized by the system and points were awarded to users as a reward when they achieved their step goals. The results showed that providing personalized step goals and rewards was effective for increased activity and sustained engagement in physical exercise. It can be as complicated as we want, but it can be as simple as allowing users to name themselves or to set their activity goals themselves as Samsung does.

6.3. Limitations

The study had also its limitations. First, the length of autoethnographic research was relatively short due to the limited time for the study. Given that the one-third of users stopped using their activity trackers within 6 months, I can identify myself as a short-term user. The use and impact of virtual rewards may change along with the evolving activity goals and practices over time. This kind of long-term use may shed light on different aspects of virtual reward experience. Second, the survey participants were predominantly younger adults (69% out of 113), which was caused by the demographic characteristics of the platform I used for recruitment. Since digital rewards, especially badges, are more likely to appeal to younger people, this composition of study population could have influenced the results obtained. Third, A relatively small number of Google Fit users participated in the study. This hinders generalizing the findings derived from these users although there was not a big difference between the results of autoethnography and their responses. Lastly, 24 Likert scale items used for quantitative research were not a complete set of tools to measure users' attitudes and perceived effects of virtual reward system in activity trackers. They were constructed by the author based on the results of autoethnography and the reviewed literature and used without iterative development. Therefore, they need to be validated and further improved before carrying out a similar study with a larger population. These limitations should be addressed in future work to further investigate the changing relationship between users and virtual reward systems, the difference in perception and usage pattern between younger and older adults, and the difference in influence and role between points and badge systems.

07

Conclusion

Conclusion

This study has investigated real-life experiences with virtual rewards provided by activity tracking systems and their effects and value in people's physical activity practices using mixed methods: a diary-based autoethnography with three different virtual reward systems of current activity trackers and an online survey with 113 participants. While answering the research questions, the study has made several academic contributions. First, it provides a rich understanding of how virtual rewards interact with users and other system features in real-life contexts, in which the elements that influence the quality of virtual reward experience was also revealed. Second, it gives further evidence on the influence and role of virtual rewards in real-life practices and where the value of virtual rewards comes from in the context of activity tracking. Finally, it offers a set of considerations for the design of virtual reward systems that can lead to more user-centric, multi-layered, and meaningful virtual reward experience. The study recommends future research that includes long-term autoethnographic research and a wider range of participants in terms of age and device of use.

References

References

- Antin, J., & Churchill, E. F. (2011). Badges in social media: A social psychological perspective. In *CHI 2011 Gamification Workshop Proceedings* (pp. 1-4).
- Berkovsky, S., Coombe, M., Freyne, J., Bhandari, D., & Baghaei, N. (2010). Physical activity motivating games: virtual rewards for real activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 243-252). ACM. <https://doi.org/10.1145/1753326.1753362>
- Cecchinato, M. E., Cox, A. L., & Bird, J. (2017, May). Always on (line)?: user experience of smartwatches and their role within multi-device ecologies. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3557-3568). ACM. <https://doi.org/10.1145/3025453.3025538>
- Columbia University. (n.d.). Content Analysis | Columbia University Mailman School of Public Health. Retrieved October 16, 2019, from <https://www.mailman.columbia.edu/research/population-health-methods/content-analysis>
- Consolvo, S., McDonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., ... & Smith, I. (2008, April). Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1797-1806). ACM. <https://doi.org/10.1145/1357054.1357335>
- Cugelman, B. (2013). Cugelman, B. (2013). Gamification: what it is and why it matters to digital health behavior change developers. *JMIR serious games*, 1(1), e3. <https://doi.org/10.2196/games.3139>
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627.
- Desjardins, A., & Ball, A. (2018). Revealing tensions in autobiographical design in HCI. In *proceedings of the 2018 designing interactive systems conference* (pp. 753-764). ACM. <https://doi.org/10.1145/3196709.3196781>
- Deterding, S. (2011). Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using game design elements in non-gaming contexts, a workshop at CHI*.
- Deterding, S. (2013). Skill atoms as design lenses for user-centered gameful design. In *Workshop Papers CHI2013*.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to

- gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15). ACM. <https://doi.org/10.1145/2181037.2181040>
- Difallah, D., Filatova, E., & Ipeirotis, P. (2018). Demographics and dynamics of mechanical Turk workers. In *Proceedings of the eleventh acm international conference on web search and data mining* (pp. 135-143). ACM. <https://doi.org/10.1145/3159652.3159661>
- Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: an overview. *Historical Social Research/Historische Sozialforschung*, 273–290.
- EMarketer. (2019). Wearables 2019 - Advanced Wearables Pick Up Pace as Fitness Trackers Slow. Retrieved October 16, 2019, from <https://www.emarketer.com/content/wearables-2019>
- Endeavour Partners. (2014). Inside Wearables, How the Science of Human Behaviour Change Offers the Secret to Long-Term Engagement. Retrieved October 16, 2019, from <https://medium.com/@endeavourprtnrs/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3>
- Fanfarelli, J. R., & McDaniel, R. (2015). Digital badges for deliberate practice: Designing effective badging systems for interactive communication scenarios. In *Proceedings of the 33rd Annual International Conference on the Design of Communication* (p. 49). ACM.
- Fitbit Inc. (2019). Help article: What should I know about my heart rate data? Retrieved October 16, 2019, from https://help.fitbit.com/articles/en_US/Help_article/1139
- Fogg, B. J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Fritz, T., Huang, E. M., Murphy, G. C., & Zimmermann, T. (2014). Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 487-496). ACM. <https://doi.org/10.1145/2556288.2557383>
- Garmin Ltd. (n.d.). Garmin Connect Badge Achievements | Garmin Support. Retrieved October 16, 2019, from <https://support.garmin.com/en-PH/?faq=6pECo6UIFn7ergw8kNmFu9>
- Goh, D. H. L., Pe-Than, E. P. P., & Lee, C. S. (2017). Perceptions of virtual reward systems in crowdsourcing games. *Computers in Human Behavior*, 70, 365–374. <https://doi.org/10.1016/j.chb.2017.01.006>
- Gouveia, R., Karapanos, E., & Hassenzahl, M. (2015). How do we engage with activity trackers?: a longitudinal study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1305-1316). ACM. <https://doi.org/10.1145/2741961.2742000>

doi.org/10.1145/2750858.2804290

- Hamari, J., & Eranti, V. (2011). Framework for Designing and Evaluating Game Achievements. In *Digra conference*.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work?-A Literature Review of Empirical Studies on Gamification. In *HICSS* (Vol. 14, No. 2014, pp. 3025-3034). <https://doi.org/10.1109/HICSS.2014.377>
- Harrison, D., Marshall, P., Bianchi-Berthouze, N., & Bird, J. (2015). Activity tracking: barriers, workarounds and customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 617-621). ACM. <https://doi.org/10.1145/2750858.2805832>
- HASTAC (Humanities, Arts, Science, and T. A. and C. (n.d.). Digital Badges _ HASTAC. Retrieved October 16, 2019, from <https://www.hastac.org/initiatives/digital-badges>
- Högberg, J., Hamari, J., & Wästlund, E. (2019). Gameful Experience Questionnaire (GAMEFULQUEST): an instrument for measuring the perceived gamefulness of system use. *User Modeling and User-Adapted Interaction*, 1-42.
- IDC. (2018a). IDC Forecasts Slower Growth for Wearables in 2018 Before Ramping Up Again Through 2022. Retrieved October 16, 2019, from <https://www.idc.com/getdoc.jsp?containerId=prUS44276818>
- IDC. (2018b). Worldwide Wearables Market Ticks Up 5.5% Due to Gains in Emerging Markets, says IDC. Retrieved October 16, 2019, from <https://www.idc.com/getdoc.jsp?containerId=prUS44247418>
- IDC. (2019). Wrist-Worn Wearables Maintain a Strong Growth Trajectory in Q2 2019, According to IDC. Retrieved October 16, 2019, from <https://www.idc.com/getdoc.jsp?containerId=prUS45521319>
- Jakicic, J. M., Davis, K. K., Rogers, R. J., King, W. C., Marcus, M. D., Helsel, D., ... & Belle, S. H. (2016). Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. *JAMA - Journal of the American Medical Association*, 316(11), 1161–1171. <https://doi.org/10.1001/jama.2016.12858>
- Jarrahi, M. H., Gafnowitz, N., & Shin, G. (2018). Activity trackers, prior motivation, and perceived informational and motivational affordances. *Personal and Ubiquitous Computing*, 22(2), 433–448. <https://doi.org/10.1007/s00779-017-1099-9>
- Jobe, W. (2014). No university credit, no problem? Exploring recognition of non-formal learning. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1-7). IEEE. <https://doi.org/10.1109/FIE.2014.7044389>

- Kappen, D. L., Mirza-Babaei, P., & Nacke, L. E. (2017). Gamification through the application of motivational affordances for physical activity technology. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (pp. 5-18). <https://doi.org/10.1145/3116595.3116604>
- Karapanos, E., Gouveia, R., Hassenzahl, M., & Forlizzi, J. (2016). Wellbeing in the making: peoples' experiences with wearable activity trackers. *Psychology of well-being*, 6(1), 4. <https://doi.org/10.1186/s13612-016-0042-6>
- Khot, R. A. (2013). Sweat-atoms: crafting physical objects with everyday exercise. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2701-2706). ACM.
- Klasnja, P., Consolvo, S., & Pratt, W. (2011). How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3063-3072). ACM. <https://doi.org/10.1145/1978942.1979396>
- Korinek, E. V., Phatak, S. S., Martin, C. A., Freigoun, M. T., Rivera, D. E., Adams, M. A., ... & Hekler, E. B. (2018). Adaptive step goals and rewards: a longitudinal growth model of daily steps for a smartphone-based walking intervention. *Journal of Behavioral Medicine*, 41(1), 74–86. <https://doi.org/10.1007/s10865-017-9878-3>
- Lewis, Z. H., Swartz, M. C., & Lyons, E. J. (2016). What's the point?: a review of reward systems implemented in gamification interventions. *Games for health journal*, 5(2), 93-99. <https://doi.org/10.1089/g4h.2015.0078>
- Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 557-566). ACM.
- Lin, J. J., Mamykina, L., Lindtner, S., Delajoux, G., & Strub, H. B. (2006). Fish'n'Steps: Encouraging physical activity with an interactive computer game. In *International conference on ubiquitous computing* (pp. 261-278). Springer, Berlin, Heidelberg. https://doi.org/10.1007/11853565_16
- Lucero, A. (2015). Using affinity diagrams to evaluate interactive prototypes. In *IFIP Conference on Human-Computer Interaction* (pp. 231-248). Springer, Cham.
- Lucero, A. (2018). Living without a mobile phone: an autoethnography. In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 765-776). ACM. <https://doi.org/10.1145/3196709.3196731>
- Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013). Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 extended abstracts on human factors in computing systems* (pp. 1137-1142). ACM.

- Montgomery-Downs, H. E., Insana, S. P., & Bond, J. A. (2012). Movement toward a novel activity monitoring device. *Sleep and Breathing*, 16(3), 913-917. <https://doi.org/10.1007/s11325-011-0585-y>
- Montola, M., Nummenmaa, T., Lucero, A., Boberg, M., & Korhonen, H. (2009). Applying game achievement systems to enhance user experience in a photo sharing service. In *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era* (pp. 94-97). ACM. <https://doi.org/10.1145/1621841.1621859>
- Munson, S. A., & Consolvo, S. (2012, May). Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In *2012 6th international conference on pervasive computing technologies for healthcare (PervasiveHealth) and workshops* (pp. 25-32). IEEE. <https://doi.org/10.4108/icst.pervasivehealth.2012.248691>
- Nicholson, S. (2012). A User-Centered Theoretical Framework for Meaningful Gamification. *Games+Learning+Society* 8.0.
- Novak, J. I., & Loy, J. (2018). A pilot study for utilizing additive manufacturing and responsive rewards in physical activity gamification. *Design for Health*, 2(2), 266-284. <https://doi.org/10.1080/24735132.2018.1541400>
- Orji, R., Vassileva, J., & Mandryk, R. L. (2014). Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. *User Modeling and User-Adapted Interaction*, 24(5), 453-498. <https://doi.org/10.1007/s11257-014-9149-8>
- Pearson, C. (2018). Project Spotlight: Badges - Chris Pearson - Medium. Retrieved October 25, 2019, from <https://medium.com/@pancakefeed/project-spotlight-badges-320aa50375f4>
- Reeves, S., Kuper, A., & Hodges, B. D. (2008). Qualitative research methodologies: ethnography. *Bmj*, 337, a1020.
- Reward [Def. 1]. (2019). In *Lexico powered by Oxford*. Retrieved from <https://www.lexico.com/en/definition/reward>
- Richter, G., Raban, D. R., & Rafaeli, S. (2015). Studying gamification: the effect of rewards and incentives on motivation. In *Gamification in education and business* (pp. 21-46). Springer, Cham.
- Samsung. (2019). Samsung Health: How to earn reward badges - YouTube. Retrieved October 25, 2019, from <https://www.youtube.com/watch?v=rLuhTOyHar4>
- Shields, R., & Chugh, R. (2017). Digital badges – rewards for learning? *Education and Information Technologies*, 22(4), 1817–1824. <https://doi.org/10.1007/s10639-016-9521-x>
- Shih, P. C., Han, K., Poole, E. S., Rosson, M. B., & Carroll, J. M. (2015). Use and adoption

- challenges of wearable activity trackers. *IConference 2015 Proceedings*.
- Skeuomorph [Def. 1.1]. (2019). In *Lexico powered by Oxford*. Retrieved from <https://www.lexico.com/en/definition/skeuomorph>
- Stusak, S., Tabard, A., Sauka, F., Khot, R. A., & Butz, A. (2014). Activity sculptures: Exploring the impact of physical visualizations on running activity. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 2201–2210. <https://doi.org/10.1109/TVCG.2014.2352953>
- Tseng, Y. C., Chang, H. Y., & Yen, S. W. (2018, April). The Different Effects of Motivational Messages and Monetary Incentives on Fostering Walking Behavior. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. LBW017). ACM.
- Vassileva, J. (2012). Motivating participation in social computing applications: a user modeling perspective. *User Modeling and User-Adapted Interaction*, 22(1-2), 177-201.
- Vooris, R., Blaszk, M., & Purrington, S. (2019). Understanding the wearable fitness tracker revolution. *International Journal of the Sociology of Leisure*, 1-17. <https://doi.org/10.1007/s41978-018-00022-y>
- Waltz, E. (2012). How i quantified myself. *IEEE Spectrum*, 49(9), 42–47. <https://doi.org/10.1109/MSPEC.2012.6281132>
- Wang, H., & Sun, C. T. (2011). Game reward systems: Gaming experiences and social meanings. In *DiGRA Conference* (Vol. 114).
- World Health Organization. (2010). *Global recommendations on physical activity for health*. World Health Organization.
- Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. “O’Reilly Media, Inc.”.
- Zuckerman, O., & Gal-Oz, A. (2014). Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Personal and ubiquitous computing*, 18(7), 1705-1719. <https://doi.org/10.1007/s00779-014-0783-2>